What Determines the Firm Size Distribution and Structural Integration? A Cross-Country Study

Manet Hun

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy (Ph.D.) in Economics in the Faculty of Social Sciences and Law

Department of Economics

April 2009

Word Count: 42,123

Abstract

A principal objective of this dissertation is to seek a better understanding of the factors affecting the firm size distribution, giving special attention to the small and medium enterprise sector. To this end, this dissertation uses two measures of the relative importance of the small and medium enterprise sector – the employment share of small and medium enterprises in the manufacturing sector and the relative size of the small enterprise sector versus the large enterprise sector.

Chapter 3 examines the determinants of the firm size distribution, using Bayesian model averaging to address the model uncertainty issue. Chapter 4 further investigates potential disproportionate effects of financial development and financial liberalization on small and large enterprises by employing a dynamic panel data approach. Chapter 5 uses instrumental variable methods to examine potential causal effects of the quality of institutions on the firm size distribution.

Some key findings in chapters 3, 4 and 5 include (1) the relative importance of small and medium enterprises in the economy initially rises and then falls, as income increases; (2) there is a positive association between the share of small and medium enterprises in the economies and effective and efficient institutions, although there is insufficient evidence to suggest a causal effect of institutional quality on the SME sector; (3) a positive and disproportionately larger effect of financial liberalization policies that open up a country's financial system for small enterprises compared to large enterprises; and (4) a dominating role of small firms in countries which lack easy access to international markets.

Chapter 6 searches for explanations for the cross-country variation in the degree of structural integration, or structural dualism, as proxied by the relative labour productivity in agriculture versus other sectors. Using Bayesian model averaging once again, we find that the extent of dualism has an inverted U-shape relationship with the level of income, and is negatively related to the initial level of human capital and physical capital accumulation. We also find that politically freer countries tend to be more dualistic.

Acknowledgements and Dedication

I would like to thank many people for their support in making this dissertation possible. I am especially grateful to my supervisor, Jonathan Temple, for his advice, guidance and support. His time, insights and constant encouragement throughout these years have contributed tremendously to my academic growth and been vital to the completion of this dissertation.

I am also thankful to Edmund Cannon, David Demery, Andrew Pickering and other members of the Macroeconomics, Growth and Development Group for their useful comments and suggestions on numerous occasions.

Many thanks to my fellow PhD students for making my study experience an enjoyable one. My special gratitude goes to Yongfu Huang for helping me with Scientific Word.

I would also like to express my appreciation to supporting staff at the Department of Economics for their assistance during my study at the University.

Last but not least, I would like to dedicate this dissertation to my parents and my wife whose continuous support and encouragement have been invaluable sources of motivation for me.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol. The work is original, except where indicated by special reference in the text, and no part of the dissertation has been submitted for any other academic award. Any views expressed in the dissertation are those of the author.

SIGNED: ..... **DATE:** .....

Contents

Chapter 1
Introduction

The importance of small and medium sized enterprises in terms of their contribution to economic growth via job creation and value added in the economy has been an ongoing debate among economists and policy makers alike. Some believe that small and medium sized firms are much more effective than their larger counterparts in creating jobs and contributing to economic growth and hence many government policies, especially in developing countries, have been implemented to facilitate the growth of the small firm sector. Others have argued that larger firms are more effective and thus suggest a policy response in that direction.

A fair amount of attention has also been given to the effects of different factors on firms of different sizes. Understanding how these factors interact and how they affect firms of different sizes is important because it allows economists to better understand what determines the firm size distribution in a country and helps policy makers structure appropriate and effective responses to meet their policy objectives, regardless of their views on the importance of small and large enterprises.

Motivated by this rationale, much of this dissertation seeks to gain a better understanding of the sources of cross-country variation in the distribution of firm size, with special emphasis on small and medium enterprises. Specifically, it will examine the relationship between various factors on the one hand and firm size and the firm size distribution on the other. In fact, a large part of this dissertation (chapters 3-5) is devoted to addressing this issue. In addition, we will look at factors affecting the degree of structural integration in the economy.

In order to do so, our analysis uses as the dependent variables two indicators of firm size distribution, which are the employment share of the small and medium enterprises in manufacturing sector and the relative size of the small enterprise sector versus the large enterprise sector. Chapter 3 will discuss in details various existing measures of the SME sector as well as the construction of the two dependent variables. Briefly, the first measure – the share of the small and medium enterprise sector – is constructed by Ayyagari

et al. (2003). Using available data from various official data sources including the World Bank and the OECD, they construct a cross-country measure of small and medium enterprises as the share of this sector in manufacturing sector employment. To allow for a consistent cross-country comparison, they apply a uniform definition of small and medium enterprises across different countries using official country definitions of SMEs. Their data are available for 76 countries, averaged over the 1990-1999 period.

This dissertation's new measure, the relative size of the small enterprise sector versus the large enterprise sector, is constructed based on the United Nations Industrial Development Organization (UNIDO) industrial statistics database. The classification of the small and the large enterprise sectors is based on the average number of employees per industrial establishment across countries. The relative size of the small enterprise sector is the ratio of the number of employees in small enterprise sectors to that in small enterprise sectors plus that in large enterprise sectors. There are 61 countries where necessary data are available on an annual basis and cover the 1970-1996 period. A more detailed description of the construction of this new variable will be given in chapter 3.

The UNIDO-based measure differs from the first measure in that it does not rely on any arbitrary definition of the SMEs. Another difference between the two measures is that the UNIDO-based measure offers time variation while the other measure does not. Moreover, whereas Ayyagari et al. measure considers small and medium enterprises, the UNIDO data take into account only small and large enterprises and exclude those in between. Therefore, due to these differences, the two measures may differ from one another. In fact, the correlation between them is low.

We begin with a survey of the literature relating to firm size structure in chapter 2. In particular, we will look at studies of the relative importance of small and medium enterprises versus large enterprises in promoting economic growth and development. We also discuss various theoretical work that seeks to explain the determinants of firm size and the firm size distribution.

In chapter 3 we examine the sources of cross-country variation in the distribution of firm sizes by employing a new econometric method, the Bayesian Model Averaging (BMA) method. This method has clear advantages over conventional approaches used in previous studies including the fact that it takes into account the model uncertainty issue that

is commonly faced by traditional cross-section analysis.[1] The analysis is divided into two parts. First, the share of the small and medium enterprise sector is used as the dependent variable. Then, the UNIDO-based measure of the relative size of the small enterprise sector versus the large enterprise sector is used.

Our findings suggest that the relative importance of small and medium enterprises declines over the long run. However, the relationship with income is non-linear. Specifically, there exists an inverted-U shape relationship between the share of the SME sector and income level, with the share of the SME sector in manufacturing and in the economy initially increasing as income rises at the early stage of development, but then declining once income reaches a certain level. The turning point, however, varies from one country to another.

The results also indicate a significant and positive relationship between the quality of institutions and the SME share in the economy. This is consistent with the view that better and more effective functioning of institutions helps create an environment in which smaller and medium sized firms find it easier to compete and to grow. Furthermore, regulations such as employment protection laws are found to have a disproportionately negative effect on small and medium size firms compared to larger firms.

A country's exposure to international trade, as measured by the ratio of exports to GDP, is found to have a negative and significant relationship with the SME sector share. This finding suggests that more open countries tend to have larger firms. In addition, empirical evidence suggests that firms in a geographically isolated country may find it harder and more costly to engage in international trade. As a result, their activities are likely limited to within their own border and it will be difficult for them to exploit larger external markets and to grow. In fact, we find that such a country's economy tends to be dominated by small and medium sized enterprises.

Unlike some previous studies, we do not find any robust association between either human capital or financial factors and the share of the small and medium enterprise sector in the economy. Similarly, the effects of a country's physical infrastructure and macroeconomic stability on the firm size distribution are not clear.

---

[1] More detailed discussion of the BMA method will be provided in chapter 3.

When the relative size of the small enterprise sector versus the large enterprise sector is used as the dependent variable, we find that the quality of institutions, a country's relative geographical isolation, urban population share in the total population, and the degree of de-industrialization have positive relationships with the dependent variable. On the other hand, human capital is found to have a negative and significant relationship with the share of the small enterprise sector. This is different from the finding when the SME sector share is the dependent variable. Another inconsistent finding concerns the share of exports in GDP, which exhibits a positive relationship with the UNIDO-based measure while having a negative relationship with the other measure of the SME sector. Potential sources of these inconsistencies may rest on the differences between the two dependent variables, as explained earlier, in terms of their construction and what they measure.

Chapter 4 further investigates potential disproportionate effects of financial development and financial liberalization, as proxied by the degree of capital account openness, on the small and the large enterprise sectors. Dynamic panel data methods are used to conduct the analysis. This chapter uses only the UNIDO-based measure of the relative importance of the small enterprise sector as the dependent variable because it offers the time series variation needed for the econometric exercise. If financial development and financial liberalization disproportionately benefit small enterprises, then we should see an expansion in the relative size of the sector dominated by small enterprises as a result of improvements in the level of financial development and the degree of financial liberalization.

Results indicate an unclear effect of the level of financial development, which is consistent with the finding in chapter 3. However, we find a positive effect of the degree of financial liberalization on the relative size of the small enterprise sector in the economy, which is different from that found in Chapter 3. A possible explanation for such inconsistencies rests on the different methodologies employed for these analyses. In this sense, the dynamic panel data approach, which takes into account variation of the variables of interest across time, could be more informative than the cross-section BMA approach. Hence, the finding in chapter 4 may provide a more accurate picture of the relationship between financial liberalization, in terms of capital account openness, and the relative size of the small enterprise sector.

Chapter 5 uses instrumental variable (IV) estimation methods to investigate potential causal impacts of institutions on firm size structure. *Institutions* refer to a country's political and regulatory environment which may affect the functioning of the market and firms operating in that country. The analysis will be conducted in two ways. First, a potential causal relationship between institutional measures and the small and medium enterprise share in manufacturing employment is examined. Second, potential disproportionate effects of institutions on firms of different sizes are investigated by examining the effects of institutional measures on the UNIDO-based measure of the relative share of the small enterprise sector versus the large enterprise sector.

Results from the analysis indicate that although our measure of institutions has a positive relationship with the share of the SME sector, this relationship is not robust to controlling for simultaneity bias. This suggests that while countries with good institutions generally have a large SME sector, there is insufficient evidence to support the view that institutional quality exerts a causal effect on the size of the SME sector.

While chapters 3, 4 and 5 focus solely on different-sized firms within the manufacturing sector, in chapter 6 we examine the inter-sectoral relationship between the agricultural and non-agricultural sectors, by looking at different factors affecting the degree of structural integration between these two sectors. Specifically, this chapter examines various sources of structural dualism in the economy, as proxied by relative labour productivity in agriculture versus other sectors. Similar to chapter 3, the Bayesian Model Averaging method is used to conduct our analysis here.

With regards to the determinants of the degree of structural integration, we find that the extent of structural dualism – i.e. a lack of intersectoral integration – has an inverted U-shape relationship with the level of income. Specifically, during the development process, countries are likely to become more dualistic at an earlier stage of development but become less so at a later stage of development, as incomes reach certain levels.

The degree of dualism is also found to be negatively related to the initial level of human capital, which may suggest potential lagged effects of the improvements in the quality of human resource on the extent of intersectoral integration in the economy. Likewise, empirical results indicate that physical capital accumulation helps to reduce structural dualism and thus improve structural integration in the economy in the long run. We also find

that politically freer countries tend to be more dualistic, perhaps due to a higher possibility of wage distortion produced by wage bargaining power of interest groups such as labour unions.

Finally, Chapter 7 will be the conclusion, which provides a general summary of the findings in this dissertation.

Chapter 2
SME, Firm Size and Firm Size Distribution: A Survey

This chapter presents a brief survey of the literature relating to firm size and firm size distribution. A more detailed literature review will be given in each of the following chapters.

The literature to be discussed will be divided into two separate categories, with the first category being devoted to discussion of the importance of small and medium enterprises in promoting economic growth. The second category of literature focuses on the determinants of firm size and the firm size distribution.

2.1 Small and Medium Enterprises and the Economy

In this section we discuss various studies that have been conducted on the role of small and medium enterprises in fostering economic growth and development.

It has been argued that SMEs contribute significantly to economic growth, employment generation, poverty reduction and a more equitable income distribution in developing countries. The World Bank suggests that SMEs enhance competition and entrepreneurship and hence have external benefits on economy-wide efficiency, innovation, and aggregate productivity growth (World Bank 2002, 2004).

An important characteristic of small and medium sized firms is that they may possess a flexibility that enables them to adapt quickly to changing market conditions and to withstand adverse economic shocks. This is true for a number of reasons. First, SMEs are more labour-intensive than larger firms and thus have lower capital costs associated with job creation. This gives these enterprises the ability to adapt to changing conditions quickly and cheaply. Second, SMEs also enjoy such flexibility because they do not depend heavily on infrastructure and because they can change their inputs and product lines at relatively low cost due to the nature of their typically low levels of technology (Morawetz, 1974). Third,

small enterprises in developing countries usually rely on personal savings and informal loans to finance their operation. Thus not only do these enterprises play an important role in mobilizing savings for productive investments, they can shield themselves from negative shocks that may affect the formal financial sector.

Various empirical studies do lend support to this premise. Berry and Mazumdar (2001), for instance, examine evidence from many countries in Southeast Asia and have come to the conclusion that SMEs are especially important in industries or economies that face rapidly changing market conditions such as the sharp macroeconomic downturns seen in the Asian financial crisis in 1997. Similarly, Mulhern and Steward (2003) argue that during an economic downturn such as a recession, large firms are more adversely affected compared to small and medium sized firms; and as a result the share of the latter in the economy tends to rise during such hard times. Such ability to withstand sharp macro-economic fluctuations means that SMEs can play a significant role in ensuring long-term economic stability and growth.

Another common argument in support of small and medium enterprises involves their choice of technology, which may be more suitable to local conditions and is conducive to employment creation. Unlike large enterprises which usually employ capital-intensive modes of production, small firms tend to employ more labour-intensive technologies and thus their expansion would generate more employment than large enterprises (Berry and Mazumbar, 1991).

In addition, in many developing countries the agricultural sector has not been able to generate sufficient jobs to accommodate a rapidly increasing number of new entrants to the labour market in rural areas. In this sense, SMEs, especially those located in rural areas, can offer off-farm employment opportunities which help absorb surplus labour and provide an alternative means for income generation for rural households.

Proponents of SMEs also argue that the traditional industrialization process is often associated with a tendency towards increasing concentration of industry in the hands of a few large firms, which leads to income inequalities and social disparities (Kitching, 1982). In this regard, small and medium enterprises offer an alternative avenue to development that results in a more equitable distribution of income because they have a much broader entrepreneurial base than large enterprises. Furthermore, as Mazumdar (2001) points out,

the typically large wage differential between SMEs and LEs implies that a larger share of output produced by the former leads to more of the wage bill going to workers in the lower wage groups.

Similarly, while large enterprise development is often associated with excessive concentration in major metropolitan areas, widespread SME growth could create many growth poles in small towns and rural areas, which could serve as the basis for renewed sources of growth and more equitable income distribution (Mazumdar, 2001). For these reasons, the expansion of the small and medium enterprise sector would not only contribute to high employment growth, but also lead to a more desirable social outcome.

While SME sector development has been a focus for economists and policy makers in developing countries for many years, recent development has seen it attract considerable amount of attention among policy makers in developed economies as well. This is partly due to the fact that, in some developed economies, the process of industrialization has peaked and a shift towards de-industrialization has occurred, with SMEs assuming more important roles in the economy.[2] For instance, in the Netherlands, small and medium enterprises account for at least 95 percent of total business establishments (Bijmolt and Zwart, 1994). Similarly, in other industrialized countries such as Australia, Canada, France, Germany, and Japan, SMEs are an important engine of economic growth and technological progress (Thornburg, 1993). In the United States, recent recognition of the importance of small and medium enterprises as a driving engine of growth, job creation and competitiveness in the global market has prompted the government to place strong policy emphasis on promoting these enterprises (Audretsch, 2000).

Nevertheless, these pro-SME views have not gone without challenge. In fact, several recent studies suggest just the opposite. For example, they argue that large enterprises are much better placed than their smaller counterparts to exploit economies of scale and are more easily able to undertake the fixed costs associated with research and development activities with positive productivity effects. Moreover, empirical evidence from both developing and developed countries show that LEs provide more stable employment, higher

[2] For more discussion on the re-emergence of the importance of small firms in developed economies, see for example, studies by Carisson (1989), Acs and Audretsch (1990), Dunne and Hughes (1990) and Sengenberger et al. (1990).

wages and more nonwage benefits than small firms, which have positive ramifications for poverty alleviation (Beck and Dermirguc-Kunt, 2004).

A number of studies including those of Little et al. (1987) and Snodgrass and Biggs (1996) show that, contrary to common belief, SMEs are neither more labour intensive, nor better at job creation than LEs. In a similar vein, Biggs and Shah (1998) examine evidence from Sub-Saharan Africa and find that LEs are the dominant source of employment creation in the manufacturing sector. In addition, Davis et al. (1993) find that there is no systematic relationship between net job creation and firm size.[3]

Arguments against the pro-SME policy also come from emerging empirical evidence which supports the view that firm size responds to national institutional conditions (Tambunan, 2005). For instant, Beck et al. (2002) find that countries with good banking systems or well-developed financial institutions tend to have more LEs than SMEs, because successful firms face no financial constraints and can grow to their efficient sizes.

2.2 Determinants of Firm Size and the Firm Size Distribution

This section discusses various theories on the determinants of firm size and the firm size distribution. Further empirical literature on this subject will be discussed in later chapters. The theories considered here are classified into four groups: technological, transaction cost, firm dynamics, and institutional and financial theories.

2.2.1 Technological Theories

According to the technological theory, firm size is dependent on technology-determined economies of scale and on firms' social organization.

Classical views of firm size including those of Viner (1932) and Baumol et al. (1982) suggest that efficient firm size is mainly dependent on technologically-determined economies of scale and that efficient firm size is achieved where long-run average cost is minimized.

[3] See also Brown et al. (1990), Pagano and Schivardi (2001) and Biggs (2002) among others for similar discussion on this issue.

It is argued that due to the large indivisibility associated with special-purpose machineries and other specialized inputs used, production technologies in manufacturing generally exhibit increasing returns to scale. Georgescu-Roegen (1972) suggests that when there are a number of indivisible inputs involved, economies of parallel multiplication operate until the scale of operation reaches the least common multiple of the indivisibility. Thus, as the number and the size of the indivisibility increases, the minimum efficient scale will also increase. Other things equal, the larger the economies of scale present in the production technology, the larger the efficient firm size.

Nevertheless, this view does not go without challenge. Some argue that while economies of scale can determine the minimum efficient scale of a single production unit, they alone cannot explain the efficient firm size partly because how many production units a firm should operate cannot be decided solely by production technology (You, 1995). This, thus, leads to another determining factor of firm size, which is a firm's social organization.

The social organization aspect involves treating firms as social organizations rather than just as sets of tools or machineries. In this case, costs are determined not only by the characteristics of the production technology but by the effectiveness of the organization in decision-making and implementation.

A principal aspect of the social organization view rests on the assumption of a fixed and non-replicable input such as the entrepreneurial or the managerial input. On the one hand, firm size is said to be determined by the individual entrepreneur's ability and propensity to take on risks (Knight, 1965, first published in 1921; Kihlstrom and Laffont, 1979). A greater ability and willingness of the entrepreneur to take risk may lead to the expansion of firm size, while the lack of such willingness limits the ability of firms to grow, resulting in small firm size.

On the other hand, Lucas (1978) equates a firm with a manager and suggests that an increase in production scale will lead to a decrease in management effectiveness. He assumes that there is a given distribution of managerial talents across agents. In equilibrium, there are managers and employees. The greater the talent of an agent, the greater the size of the firm, as measured by the number of employees he or she manages. A central aspect of this approach is the decision an agent faces between becoming a manager or an employee, and the decision a manager faces on the optimal choice of the levels of employment and

capital in his firm. An increase in per capita capital raises wages relative to managerial rents, inducing marginal managers to become employees and thus increases the ratio of employees to managers, resulting in larger firm size. For this reason, average firm size may be positively associated with income per capita (see Kumar, Rajan and Zingales, 1999). In addition, using time series data in the US, Lucas finds a positive relationship between firm size and the amount of capital per capita, as proxied by GNP per capita. This finding suggests that more capital intensive technologies are associated with larger firm size.

On a similar note, Marx (1976, first published in English in 1887) observed that as the technical division of labour becomes more minute and as divisible labour is replaced by large-scale indivisible machines, the size of firms gets larger. Because technologies with large economies of scale tend to be capital-intensive and huge capital outlays are usually required to exploit economies of scale, the cross-industry variation in firm size will be related to capital-intensity. Moreover, empirical studies such as those of Caves and Uekusa (1976) and Banerji (1978) show that cross-country variation in firm size is related to economy-wide capital intensity and hence the level of income.

2.2.2 Transaction Cost Theories

This section discusses transaction cost theory in which firm size is determined by transaction cost efficiency. The basic concept behind the transaction cost theory is that all transactions involve costs and that different types of governance are required for different types of transactions in order to minimize these transaction costs (Williamson, 1985).[4]

Coase (1937) views a firm as an alternative to the market as a mechanism for resource allocation. He points out that within a firm, market transactions are eliminated and the complicated market structure with exchange transactions is substituted by the entrepreneur-coordinator who directs production. Therefore, firm size may be said to depend on the number of transactions the entrepreneur organizes.

For Coase, the main reason to establish a firm is to achieve lower costs than the market by avoiding some of the transaction costs of using the price mechanism.[5] It follows

[4] See also Alchian and Demsetz (1972) and Bowles and Gintis (1990) among others for discussion on the sources of transaction costs.

[5] These costs include the costs of discovering relevant prices, as well as the costs of negotiating and writing

that the maximum size of the firm is determined where the extra cost of internalizing one more transaction equals the market transaction cost. On the other hand, diminishing returns to management seems to be the principal contributor to the rising costs of organizing a large firm, which can limit the maximum size of a firm.

For these reasons, Coase concludes that the size of the firm is dependent on the costs of using the price mechanism and on the costs of organization of other entrepreneurs. These two factors together influence the entrepreneur's decision on how many products to produce and how much of each.

Other aspects of the transaction cost theory of the firm concern governance costs and asset specificity issues. With regards to the governance cost issue, the transaction cost theory implies that the efficient firm size increases when organizational innovations reduce bureaucratic costs.

The transition from the unitary organizational form, U-form, to the multidivisional organizational form, M-form, which resulted from pressures of horizontal expansion by firms serves as a good example of such innovation (Chandler, 1966). Williamson (1985) argues that the M-form serves to economize on bounded rationality as well as opportunism, thereby aiding corporate expansion. Another example of the influence of governance costs on firm size is the case where rigidities and conflicts in labour relations lead large firms to vertically disintegrate or otherwise trim down the size of the workforce (You, 1995).

Another issue concerns asset specificity. Usually, asset specificity is defined as the extent to which the investments made to support a particular transaction have a higher value to that transaction than they would have if they were redeployed for any other purpose (McGuinness, 1994). Examples of such transaction-specific assets include non-redeployable physical and human investments that are specialized and unique to a task. Williamson (1985) points out that as the degree of asset specificity and the frequency of interfirm transactions increase, efficient governance changes from market governance to contractual governance (interfirm co-operation) to unified governance (integration). In other words, the presence of an asset specificity problem may be associated with large firm size because firms seek to minimize their costs by centralizing the production of all of their products,

---

enforceable contracts for each exchange transaction, which can be large if there is uncertainty.

usually through integration.[6] In contrast, as Milgrom and Roberts (1990) argue, decreases in asset specificity due to the introduction of flexible manufacturing technology may have caused vertical disintegration in manufacturing in recent periods.

While integration has been a strategy adopted by large firms in many countries as a means to reduce transaction costs, it is not the only solution to the problem. In fact, in some countries such as Korea, Japan and Taiwan, interfirm cooperation is preferred to integration.[7] In these countries, firms rely more on subcontracting networks for supply and intermediate products.

A principal motivation for firms to establish such cooperation is that they may be able to reap some of the benefits of integration without incurring the costs of integration such as bureaucratic inflexibility (Lazonick, 1990). A well-functioning cooperative network of small firms can also have a competitive advantage over large integrated firms owing to the benefits of flexibility and specialization (Piore and Sabel, 1984) and informational efficiency (Aoki, 1988). All of these work to reduce costs of interfirm transactions.

For such cooperative arrangements to work, however, there must be ways to deal with problems which cause integration in the first place. In this sense, cultural and ideological factors may play a vital role in making such interfirm cooperation work well because these factors work to mitigate problems such as opportunism and lack of trust between parties or firms involved in the transaction. As Sengenberger and Pyke (1990) suggest, the existence of such institutions and ideologies that are capable of sustaining cooperative relations is an important factor favouring interfirm cooperation over integration For these reasons,

---

[6] According to Williamson (1975), the existence of asset specificity in production may cause problems if the assets are owned by different firms because it will lead to protracted bargaining concerning the gains from trade. As a result, both agents are likely to become locked into a position where they are competing with only a small number of agents in the market. If the transaction is a recurring or lengthy one, re-negotiation may be necessary as a continual power struggle takes place concerning the gains from trade, causing transactions costs to rise. In addition, opportunism and lack of trusts may also contribute to rising transaction costs. There may be situations where a buyer may seek to re-negotiate the purchase contract of a product after it has been made. In this case, the seller has an enormous disadvantage because he has already incurred costs for producing this product which cannot be easily sold to another buyer or used for another purpose. An effective way to address this problem may be the removal of one of the parties, e.g. seller, from the equation by takeover or merger. Another way is for the buyer to control their own source of inputs by setting up subsidiaries to ensure a stable supply of inputs and ensure a consistent quality in their final product. This process calls for the centralization of production. Both of these approaches lead to an increase in firm size.

[7] Subcontracting networks also exist in some European countries such as Italy (see Brusco, 1982; Goodman and Bamford, 1989; Pyke et al., 1990 for discussion on this).

we can conclude that cooperative arrangements between firms such as a subcontracting system would lower the extent of integration and strengthen the small firm sector. In fact, empirical evidence shows that there are more flourishing small firms in countries where interfirm cooperation is more widespread.[8]

2.2.3 Dynamic Theories

While in both technological and transaction cost theories firm size is said to be dependent on efficiency considerations, dynamic theories focus on the relationship between firm characteristics, such as age and size, and growth rates.

An important aspect of the dynamic theory of the firm concerns the age-growth relationship. One of the most influential works on this subject is the Marshall (1948, first published in 1890) life-cycle of firms which puts a strong emphasis on the importance of firm age as a determining factor for its growth rate. The average growth rate of firms is said to have an inverse relationship with age, with older firms experiencing slower but more stable rates of growth. In contrast, new firms tend to experience larger fluctuations in growth rates and have a higher failure rate compared to their older counterparts (Dunne et al., 1989a).

Jovanovic (1982) develops a model of the life-cycle based on learning. This model assumes that firms have different efficiencies which can only be obtained gradually and based on the information it receives from carrying out its operation over a long period of time. One implication of this model is that as a firm ages, the probability for it to fail decreases because older firms have more precise estimates of their true efficiencies, and therefore it is less likely that these firms will be surprised and revise their estimates steeply in the next period. This implies lower failure rates for older firms (You, 1995). Moreover, learning not only increases the chance of firms to survive, but also allows them the opportunity to grow and expand.[9] For this reason, we can infer that there is a positive correlation between firm age and size. We can also say that the firm size distribution is

[8] See, for instance, Taymaz and Kaliçaslan (2005) for empirical evidence on the relationship between subcontrating arrangements and the performance of SMEs in Turkey.

[9] This is consistent with the view of Pakes and Ericson (1987) who suggest that firms normally enter an industry at a sub-optimal scale in order to learn and expand subsequently if successful.

likely to be skewed towards small firms when there is a large presence of new firms in the industry.

Another important aspect of the dynamic theory of firms concerns the size-growth relationship. Gibrat (1931) seeks a model of firm and industry dynamics to explain his finding that the firm size distribution within an industry tends to be highly skewed and approximately lognormal. His explanation, which has become known as Gibrat's law of proportionate growth, suggests that the size of a firm and its growth rate are independent. It follows that the firm size distribution is purely a result of a stochastic growth process. In other words, as Simon and Bonini (1958) point out, chance rather than economies of scale, transaction costs, or any other systematic factors is the explaining factor for the size distribution of firms. A number of empirical studies including that of Audretsch et al. (2002) also support this growth-size independence hypothesis.

However, many scholars have questioned both the assumption and the implication of Gibrat's law. In his studies of firms across a number of industries, Mansfield (1962) finds a negative correlation between firm size and growth, which thus violates the assumption of growth-size independence of Gibrat's law. In particular, small firms are generally found to experience a higher growth rate than large firms. More recent empirical studies by Hall (1987), Evans (1987) and Dunne et al. (1989b) result in a similar finding.[10]

Another aspect of Gibrat's law that has come under criticism is the implication that the variance of the firm size distribution will increase continuously. However, as discussed earlier, the firm growth rate tends to be negatively correlated with both age and size of the firm. As a result, the variance of the firm size distribution is likely to decrease in the long run as firm sizes converge.

2.2.4 Institutional and Financial Theories

This section discusses some previous studies of the relationship between institutional and financial factors on the one hand and the firm size distribution, on the other. More discussion of these issues will be provided in later chapters.

---

[10] Further surveys by Scherer (1980), Geroski (1995), Sutton (1997) and Caves (1998) also yield similar results.

There are several channels through which institutional and financial factors can affect firm size and firm size distribution beyond that predicted by the technological, transaction cost and dynamic theories of the firm.

Political and regulatory institutions can affect firm size and the firm size distribution in a number of ways. The application of certain regulations can tilt the playing field in favour of either small firm or large firm sectors, depending on the relative cost imposed on each sector by these regulations. For instance, high corporate tax rates which impose a huge burden on large firms would drive many economic activities into the informal sector and thus tilt the firm size distribution towards small firms. Pagano, Panetta and Zingales (1998) argue that this is the reason why only a limited number of firms in Italy go public. This, in turn, restricts access to the public equity market and may limit the size of firms.

Similarly, high barriers to entry resulting from specific regulations (or their absence) could have a large effect on average firm size in a country because it could make it prohibitively costly to start up a firm and thus favour large firms. Hopenhayn (1992) develops a dynamic model of firm size based on entry costs and firm-level productivity shocks. However, his results were ambiguous. He finds that, on the one hand, output price increases with the entry cost, leading to higher employment. On the other hand, the threshold productivity level at which firms exit decreases, which increases the fraction of firms with lower employment or smaller firms.

Financially, a principal obstacle to firm growth is the availability of external finance. Moreover, small firms are known to have the most disadvantages in the credit market. Hughes (1992) examines evidence in the UK and finds that small firms usually have higher financing costs than large firms, and their growth may be constrained by inadequate availability of finance or the terms on which it is available. There is also evidence that, in periods of tight money, bank lending to small firms contracts more than that to large firms (Gertler and Gilchrist, 1991). In addition, Cabral and Mata (2003) show that, during the first stages of development, the presence of financial constraints causes the firm size distribution to be significantly right-skewed with a large mass of small firms. In future periods, when financing constraints are relaxed, some of the small firms will grow to their optimal size and thus give rise to a more symmetric distribution of firm size. For these reasons, we should expect a positive relationship between the level of financial development and firm size. Financial

development which helps ease such financial constraints would serve as a great facilitator for firm growth, especially for small firms.

In a similar vein, financial repression due to policy distortions can also have a disproportionately adverse effect on small firms. For example, the government-induced segmentation of credit markets through preferential loans to large firms is often considered as an important source of disadvantage for small firms. Thus, financial liberalization that eliminates such distortions and introduces greater efficiency and competition in the credit market would be beneficial for small enterprises. Laeven (2000) shows that small firms gain most from liberalization, because the favoritism of preferential credit directed to large firms tends to disappear under liberalization. He finds that liberalization affects small and large firms differently. Small firms are financially constrained before liberalization begins but become less so after liberalization. On the contrary, the financing constraints on large firms are low both before and after liberalization, suggesting that financial liberalization has no significant effect on large firms.

2.3 Conclusion

This chapter shows the continuous debate on the importance of small and medium enterprises as the engine for economic growth, employment generation and poverty reduction. Despite the fact that interest in SMEs has grown recently to include some developed countries, there is still no consensus among scholars that SMEs are better than large enterprises in promoting growth. In fact, empirical studies in some countries have found evidence to the contrary.

In this chapter, we also covered some theoretical discussion on the determinants of firm size and the firm size distribution. Four groups of theories – technological, transaction cost, firm dynamics and institutional and financial theories – are discussed. While each theory provides useful insights on potential effects of a number of factors on firm size and firm size distribution, none provides a complete account of all possible determinants. Further empirical investigation is thus needed to seek out other determinants of firm size and the firm size distribution, especially factors that cannot be explained by existing theories.

Chapter 3
Determinants of the Firm Size Distribution

This chapter seeks to complement a number of studies on firm size structure by examining various determinants of the firm size distribution using a fresh empirical method. The Bayesian Model Averaging (BMA) method is used in our cross-country study because it enables us to conduct an empirical analysis while addressing various issues, notably the model uncertainty problem, which are commonly present in conventional empirical research. The investigation is broken into two parts with a slight difference in emphasis. In the first part we examine various possible determinants of firm size structure by conducting the analysis with the employment share of the small and medium enterprise sector as the dependent variable. The second part will use an alternative measure of the relative importance of the small enterprise sector as the dependent variable.

Our findings suggest that the relative importance of small and medium enterprises declines over the long run. However, the relationship with income is non-linear. Specifically, there exists an inverted-U shape relationship between the share of the SME sector and the income level, with the share of the SME sector in manufacturing and in the economy initially increasing as income rises at the early stage of development, but begining to decline once income reaches a certain level. The turning point, however, varies from one country to another.

The results also indicate a significant and positive relationship between the quality of institutions and the SME share in the economy. This is consistent with the view that better and more effective functioning of institutions helps to foster an environment in which smaller firms find it easier to compete and to grow. Furthermore, regulations such as employment protection laws are found to have a disproportionately negative effect on small and medium size firms compared to larger firms.

The relationship with a country's exposure to international trade is rather ambiguous. On the one hand, the share of exports to GDP is found to have a negative and significant relationship with the SME sector employment share, suggesting that more open countries tend to have larger firms. On the other hand, the relationship becomes positive when the

UNIDO-based measure is used as the dependent variable. The latter finding implies that a more open trading system may benefit smaller size firms more than their larger counterparts. A possible explanation for this inconsistency rests on the differences between the two measures of the small enterprise sector in terms of their construction. Other potential sources of inconsistency will be discussed in a later section of this chapter.

Unlike some previous studies, we do not find any robust association between either human capital or financial factors and the share of the small and medium enterprise sector in the economy. Similarly, the effects of a country's physical infrastructure and macroeconomic stability on the firm size distribution are not clear.

One contribution of this study to the existing literature on this subject is the fact that we take into account a wider range of economic, financial, institutional and geographical indicators than most previous studies. A second important contribution is the application of a new empirical method, namely Bayesian Model Averaging, in exploring the determinants of firm size structure. The use of Bayesian Model Averaging enables us to determine robustness across a wide range of model specifications and thus differs from the conventional approach used in cross-country analysis, where results from a small set of preferred models are presented and tests for robustness using various controls is performed. Research using the conventional approach usually presents different models and it is often difficult to reconcile their results (Malik and Temple, 2005). For these reasons, a formal approach to model uncertainty and model selection such as the BMA method is needed.

In Section 2 we will discuss possible theoretical and empirical explanations of the relationship between a number of potential determinant variables and the distribution of firm size. Section 3 is a literature review where we will discuss some of the previous studies on the determinants of firm size. Section 4 will describe different variables used in our analysis. A description of the empirical framework and methodology forms Section 6, while the results will be presented in Section 7. Finally, Section 8 will be the conclusion.

3.1 Factors Affecting the Firm Size Distribution

In this section we will discuss potential relationships between the firm size distribution and a number of factors including economic, financial and institutional factors.

3.1.1 Income and Firm Size Distribution

Conventional wisdom holds that the share of small and medium enterprises in the economy decreases as the country becomes richer. One possible theoretical explanation is that such an evolution in firm size structure may be due to secular changes in the firms' output composition over time. As Weeks (2003) points out, this hypothesis maintains that at the early stage of development the economy tends to be dominated by small firms engaging in the production of basic consumer goods. As income rises, however, the production share of these consumer products falls, leading to a decline in the relative contribution of small firms to employment and output.

Another plausible explanation for this upwards move in firm size concerns the effect of economies of scale. This view maintains that the development process involves the transformation from kinship and patronage based family establishments to a capitalist method of business management, which may be more responsive to market incentives. Associated with this emerging form of management are economies of scale, perhaps conforming to the standard neoclassical presentation of long-run average cost curves. Since the internal markets of developing countries are initially small, the expansion of those markets from a low-income base would tend to induce firms to move down their long-run average cost curves to a larger scale of production (Weeks, 2003).

On this same note, Anderson (1982) argues that the evolution of the firm size distribution involves distinct stages, beginning with small-scale household and artisan production being replaced by small workshops with wage labour in the earliest stage of development. As an economy develops, these small workshops then decline in favour of medium-sized units which, in turn, are replaced by large-scale units.

A number of empirical studies also lend support to this particular view. Little, Mazumdar and Page (1987) examine empirical evidence from manufacturing firms in a number of developing countries and conclude that there is a general fall in the very small

manufacturing establishments – i.e. those with fewer than 10 workers – and they are even destined to near extinction unless they become a protected species. They further suggest that manufacturing establishments with 10-50 workers will decline relatively and probably also in absolute numbers, though the future of the mid-sized manufacturing establishments, say 50-500 workers, is harder to predict.

Mulhern and Steward (2003) examine evidence from the manufacturing sector in Venezuela and find that the SME share in the country has declined over the long run. They attribute such a decline to the failure of the small firms to improve their efficiencies and innovation, which limits their growth potential and, for some firms, even their ability to survive.

Nevertheless, this view does not go without challenge. Kumar, Rajan, Zingales (1999) examine empirical evidence from a number of European countries and do not find any robust relationship between income level and firm size. In particular, they do not find any strong evidence to support the view that richer countries have larger firms.

Furthermore, Weeks (2003) examines evidence across 25 countries and finds that there exists a non-linear relationship between income level and the share of the SME sector, as measured by the employment share of this sector in manufacturing. The SME share initially declines as income rises, but as income reaches a certain level the trend reverses. He argues that at a later stage of development, improvements in financial market and human resources help improve small firms' abilities to survive and grow, while other factors such as outsourcing help tilt the playing field away from large firms. Similarly, Ayyagari, Beck and Demirguc-Kunt (2003) investigate possible linkages between a number of factors and the share of the SME sector using a newly created data set for the share of the SME sector for a number of developed and developing countries. Their findings suggest that countries with higher GDP per capita have larger SME sectors in terms of their contribution to total employment and GDP. Other empirical investigations in some industrial economies also show that such a revival of the share of small enterprises in manufacturing and the total economy can take place at some point in the development process, though the turning point may vary from country to country.[11]

---

[11] These studies examine evidence from a number of developed nations including the US (Birch, 1981; Acs and Audretsch, 1990), the UK (Storey, 1994; Doi and Cowling, 1998), and other European countries

There is also a common structural change in many developed economies towards de-industrialization and expansion of the service sector at a later stage of development, which results in a decrease in average firm size and a rise in the share of small enterprises in the economy. The primary reason may be that the service sector is generally characterized by lower scale economies and the demand for services tends to be more customized, dedicated and specialized (You, 1995).

Such a lack of coherent findings regarding the relationship between firm size structure as indicated by the share of the SME sector and the level of income gives support to further investigation.

3.1.2 Human Capital and Firm Size Distribution

As in the case of income, there is no consensus among researchers on the effects of the availability and quality of human resources on firm size. A higher level of human capital in a country may either promote larger firms, due to the availability of higher managerial skills, or more and thus smaller firms, due to widely available entrepreneurial skills.

Kremer (1993) suggests that there is a positive correlation between human capital and firm size, arguing that a greater availability of qualified workers makes it possible for firms to specialize in more complex goods and acquire technologies that demand larger and more complicated production processes. This leads to an expansion in firm size.

On the contrary, Lucas (1978) argues that a shortage of critical human resources such as management talent may bring about the organization of larger productive units in order to better utilize this scarce resource. In a similar way, one can argue that the availability of workers with higher levels of education makes smaller scale production possible because these workers may be better in taking responsibility for more creative activities, which require greater motivation and are carried out more effectively in smaller units (Brynjolfsson, 1994).

In addition, a number of studies including Kumar et al. (1999) find no evidence to support the hypothesis that countries with higher average human capital have larger firms.

(Loveman et al., 1990, Loveman and Sengenberger, 1991) such as Italy (Trau, 1997), Greece (Thomasdakis and Droucopolous, 1996) and Norway (Spilling, 1998). Further evidence from the Far East comes from Korea (Nugent, 1996) and Taiwan (Ming-Wen Hu, 1999).

Therefore, further empirical examination of the relationship between human capital and the firm size distribution is warranted.

3.1.3 Finance and Firm Size Distribution

In this section we will discuss the possible relationship between the distribution of firm size on the one hand and the level of financial development, as well as the degree of financial liberalization (or financial openness), on the other.

It is a well known phenomenon that financial constraints can represent one of the biggest obstacles to firms' growth. According to the World Business Environment Survey conducted by the World Bank and the Inter-American Development Bank, businesses in Latin America single out the lack of external finance as the most severe obstacle to business development.

This type of constraint is most likely to affect small firms due to their general lack of abilities to mobilize sufficient internal capital needed to grow and expand. Moreover, several studies including that of Fazzari et al.(1988) and Evans and Jovanovic (1989) have shown that financial constraints are a significant determinant of firms' investment decisions, especially for young firms. Nabli and Nugent (1992) examine cross-country evidence from 54 developed and developing countries and find that development of the credit market shifts the distribution in favour of small and medium enterprises.[12]

Furthermore, using a comprehensive data set of Portuguese manufacturing firms, Mata and Cabral (2003) show that, during the first stages of development, the presence of financial constraints causes the firm size distribution to be significantly right-skewed with a large mass of small firms. In future periods, when financing constraints are relaxed, some of the small firms will grow to their optimal size and thus give rise to a less skewed distribution of firm size.

Another aspect of finance that may affect the firm size distribution is the degree of openness (or liberalization) of the financial system. McKinnon (1973) and Shaw (1973) hypothesize that financial repression due to government policies – such as interest rate

[12] On the other hand, they find that equity market development tends to favour large firms because the finance available from equity markets goes almost exclusively to large enterprises.

controls, lending rate ceilings, credit allocation at below the market clearing rate, control on international capital flows – suppresses development in the financial sector. It follows that the removal of such distortionary policies will increase both the quality and quantity of financial resources available to firms. Henrekson (1992) examines evidence from Sweden where the credit market was tightly regulated until the late 1980s, and finds that the credit regulations clearly favour credit access by larger, older, firmly established firms and by capital-intensive firms with ready sources of collateral.

Financial liberalization can introduce greater efficiency and competition in the credit market resulting in lower lending rates, which is beneficial for small enterprises. Some theories imply that financial liberalization primarily benefits small firms because they are able to respond more quickly to changing economic conditions than their larger counterparts. Small firms can respond more flexibly under difficult and changing conditions because they do not depend heavily on infrastructure, and because they can change their inputs and product lines at relatively low cost due to the nature of their typically low levels of technology (Morawetz, 1974).

Laeven (2000) argues that financial liberalization reduces imperfections in financial markets by reducing the agency costs of financial leverage. Small firms gain most from liberalization, because the favoritism of preferential credit directed to large firms tends to disappear under liberalization. He finds that liberalization affects small and large firms differently. Small firms are financially constrained before liberalization begins but become less so after liberalization. The financing constraints on large firms, however, are low both before and after liberalization. The initial difference between small and large firms disappears over time. Laeven hypothesizes that financial liberalization has little effect on the financing constraints of large firms because they have better access to preferential directed credit in the period before liberalization. This means that liberalization of the financial system would tilt the playing field in favour of small firms.

3.1.4 Institutions and Firm Size Distribution

Political and regulatory institutions can have important effects on the firm size distribution because they can influence business environments, which may either be conducive or ham-

pering to growth of firms of different sizes. Institutional efficiency and effectiveness can work to relax constraints on firms by leveling the playing field for firms of all sizes. Kumar et al. (1999) find that countries with better institutions, as measured by judicial system efficiency, tend to have a lower dispersion in firm size within an industry. They also find a positive correlation between the quality of institutions and firm size.

In addition, the application of certain regulations can affect average firm size in various ways. For example, regulations that create barriers to entry and endorse monopolies would increase average firm size, whereas such regulation as anti-trust laws would limit firm size.

A number of empirical studies have been undertaken to examine the effects of regulatory institutions on firm size structure. Davis and Henrekson (1997) argue that national differences in institutions and economic policies are the primary determinants of cross-country differences in the size distribution of firms. They support this claim by conducting a detailed study of US-Swedish differences at a single point in time. They find strong evidence which suggests that Swedish policies strongly disfavored smaller firms, entry by new firms and individual and family-owned businesses. Henrekson and Johansson (1999) investigate this relationship further by looking at the evolution of the firm size distribution over the period between 1968 and 1993. Their analysis of the institutions and rules of the game which determine the entrepreneurial and business conditions in Sweden indicates that the conditions have been unfavorable for small firms. As a result, too few small firms have managed to grow out of the smallest size classes. These findings explain why the Swedish economy is dominated by large firms. Using a more comprehensive database from across 29 industrial sectors in 15 OECD countries, Fonseca and Utrero (2006) also find that institutional factors such as labour regulations and barriers to entrepreneurship are important determinants of cross-country differences in firm size.

Schivardi and Torrini (2003, 2004) study the effect of employment protection legislation in Italy on the firm size distribution, focusing specifically on the legislation governing dismissals of workers. They hypothesize that this employment protection legislation imposes higher costs of firing workers for firms with size exceeding a certain threshold.[13]

[13] In their case, the size threshold is determined in terms of the number of employees which should not exceed 15.

Crossing this threshold would imply a higher potential cost to firms, which may discourage firms to do so and thus reduce average firm size. However, their empirical analysis shows that employment protection legislation exerts only a modest influence on the firm size distribution and, thus, it is unlikely to be a major source of cross-country differences in firm size. Therefore, whether regulatory burdens on firms increase or decrease average firm size is an empirical question.

3.1.5 Infrastructure and Firm Size Distribution

Studies including those of Easterly and Rebelo (1993) and Canning and Pedroni (1999) have found that physical infrastructure exerts a significant effect on aggregate productivity and growth.[14] Nabli and Nugent (1992) argue that poor physical infrastructure results in higher transportation costs which, in turn, exert a negative effect on average firm size because higher transport costs may cause some establishments to trade-off scale economies for lower transport costs. In other words, some firms may choose to stay small rather than having to bear extra costs that they may face if they choose to engage in external trading activities.

Similarly, Clark, Dollar and Micco (2004) and Micco and Serebrisky (2004) examine the effect of transport costs and infrastructure on trade and find that high transport costs have a disproportionately larger effect on trade for developing countries like those in Latin America than for more developed countries like the US.

Therefore, poor physical infrastructure is likely to be associated with smaller firm size as such a lack of good infrastructure may prevent firms from fully exploiting economies of scale and thus remain small. In addition, high transportation costs also reduce the exposure of firms to international markets and thus limit their ability to grow, particularly for those firms involved in the production of tradable goods.

Nevertheless, as Herrera and Lora (2005) point out, one can argue that lack of good infrastructure may have a positive effect on firm size because it may induce certain firms to operate on a larger scale, and with higher levels of vertical integration, in order to be

[14] Studies done by Canning, Fay and Perotti (1994), Fay and Perotti (1994), and Sanchez-Robles (1998) have also come to a similar conclusion.

able to internalize the supply of infrastructure. For these reasons, it is of interest to conduct further empirical investigation on the relationship between physical infrastructure and firm size.

3.1.6 Trade and Firm Size Distribution

International trade may be an important determinant of firm size structure. Exposure to international trade gives firms the ability to exploit potential economies of scale, the opportunity to raise funds at a lower cost, benefits from bulk purchasing, and a higher capacity for taking risks, such as development of new products, due to internal diversification. Wagner (1995, 2001) shows that there exists a positive relationship between firm size and direct export activities after controlling for other relevant characteristics of firms including human capital intensity, technology and innovation. This line of reasoning implies that countries with more open trade regimes will tend to have larger firms.

3.2 Previous Studies

There exist a number of studies, both theoretical and empirical, on the determinants of firm size structure. One such study, which was also discussed in chapter 2, is the You (1995) survey of different theories of the determinants of firm size and the distribution of firm sizes. He examines different theories of the firm size distribution including technological, transaction cost, industrial organization and dynamic firm theories. He concludes that, because of the highly heterogeneous nature of small firms, one cannot expect a single theory or approach to explain everything. Which theory is most suitable in explaining firm size distribution is dependent on the context. For instance, while technological theories are clearly applicable in explaining cross-industry differences and differences between rich and poor countries in average firm size, transaction cost theories are best at explaining cross-country differences in the size distribution of firms among countries of similar income levels.

Kumar et al. (1999) also examine various theories on the determinants of firm size across industries and across countries, but proceed to test these theories using empirical

evidence from a number of European countries. They find that, at the industry level, firms facing larger domestic markets are larger; while, at the country level, average firm size tends to be positively correlated with the level of development of the financial market and judicial efficiency. While this study provides us with some useful insights on the relationship between a number of factors and firm size structure, it is based on a very small sample of developed European countries. Therefore, any attempt to generalize these findings must proceed with caution. In the light of such limitations, Ayyagari et al. (2003) and Herrera and Lora (2005) conduct cross-sectional analysis of the shares of small and medium enterprises and of the share of the largest enterprises, respectively, on a number of economic, financial and institutional variables using larger samples of countries.[15] Their larger sample sizes provide more representative findings on the relationship between firm size and different factors of interest.

Our analysis differs from the above studies in a number of ways. First, it differs from the You (1995) theoretical survey in that while our study draws on theoretical explanations of the relationship between certain factors and firm size distribution, our main analytical approach is empirical. Second, this study employs a larger set of sample countries, comprising both developed and developing countries, than the one used by Kumar et al. (1999). Third, it consists of pure cross-country data which is different from Herrera and Lora (2005) dataset, which relies on firm-level data. Finally, unlike many previous studies including Ayyagari et al. (2003) and Herrera and Lora (2005), this study employs Bayesian Model Averaging (BMA) which takes into account a number of the problems commonly associated with the conventional cross-sectional approach. More detailed discussion of the BMA approach will be given in a later section.

3.3 Sample Countries and Data

This section will provide details on the variables used and their respective sources. Specifically, it will begin with the discussion of the dependent variables used in this study, giving

[15] Ayyagari et al. (2003) use pure cross-country data from 75 different countries, while the Herrera and Lora (2005) dataset draws on firm-level data (around 22,000 publicly traded companies) from 46 developed and devloping countries.

special attention to discussion of existing measures of the small and medium enterprise sector. It will also discuss the construction of a new dependent variable, which is the relative size of the small enterprise sector versus the large enterprise sector. Subsequently, a discussion of various candidate explanatory variables and their respective sources is provided.

3.3.1 Dependent Variable

Existing Measures of the Small Enterprise Sector

Over the years a number of measures of the relative importance of small and medium enterprises have been used by scholars as well as government and international organizations such as the Inter-American Development Bank (IADB), the Organization for Economic Cooperation and Development (OECD), the Asia-Pacific Economic Cooperation (APEC), the World Bank and the United Nations. The Inter-American Development Bank's SME Observatory database has time series observations on SME size and activity in a number of countries in Latin America. The OECD's Globalization and SMEs database compiles information on small and medium enterprise sectors from OECD countries, whereas the APEC's Surveys on the SMEs covers a number of countries in APEC. Similarly, while the World Bank's Regional Program on Enterprise Development has data on the structure of labour markets which also contains statistics on the SME contribution to employment in Africa, the United Nations' UN-ECE produces annual statistics and trends in national SME development focusing only on countries in transition.

The drawback of these official sources of data, however, is that there is a potential lack of coherence among them in terms of their definition and their classification of small and medium enterprise sectors as well as their coverage over countries and time. Each dataset is calculated based on their respective organization's definition and classification of small and medium size enterprises, which usually varies from one organization to another. In addition, each database covers only a number of countries within a certain region or an economic group. They generally cover a short time frame, mainly in the 1990s.

Based on data from various official data sources including those mentioned above, Ayyagari, Beck and Demirguc-Kunt (2003) constructed a new cross-country database on small and medium enterprises. This new dataset addresses the issue of a lack of coherence

in terms of defining and classifying the SME sector by taking 250 employees as a common cutoff for the definition of an SME – they call it SME250. Ayyagari, Beck and Demirguc-Kunt also construct another measure of small and medium enterprise size using official country definition of SMEs. This new measure is denoted as SMEOFF, and is the share of the SME sector in total official labour force when the official country definition of SMEs is used, with the official country definition varying between 100 and 500 employees. The two measures – SME250 and SMEOFF – are highly correlated. For this dissertation, SMEOFF is used as a dependent variable because it contains a larger number of observations.

One limitation of this dataset is that it covers only a relatively short period of time – between 1990 and 1999 – which makes it difficult to examine changes in the relative importance of small and medium enterprises over time, including responses to policy changes and financial development.

A New Measure of the Relative Size of the Small Enterprise Sector

In this chapter, we construct a new measure of the relative size of small enterprises based on the number of manufacturing establishments and their respective employment drawn from the United Nations Industrial Development Organization (UNIDO) database. The United Nations Industrial Development Organization's industrial statistics database (2005), at the 3-digit level of ISIC code, is classified into 29 different establishment categories and there are 65 countries where data are available. In addition, we fill in any missing sectoral data using a linear interpolation method based on nearby years for which data are available. Linear interpolation is a simple form of interpolation methods and it is basically a method for constructing new data points between two existing (known) data points in a linear fashion. In our study, for a particular country, linear interpolation method is used to generate missing value(s) of average number of employees per establishment for a sector and in a particular year(s) based on values of average number of employees per establishment in that particular sector in other years where data are available.

The basic rationale for the new indicator is to create an objective measure of the relative size of the small enterprise sector using available industrial data, which helps to avoid using an arbitrary definition of firm size that is normally used in the construction of other measures of small and medium enterprises, as described above.

The procedure involves a few simple calculations and is divided into two stages, which will be discussed briefly here. Appendix 3.1 provides more detailed illustration of the procedure. The first stage is the sectoral classification, which describes how UNIDO's establishment categories can be divided into small and large enterprise sectors. First, using the values for 1985,[16] the average number of employees per establishment in each category are calculated and then the respective median values across countries for these average number of employees are obtained.[17] These median values indicate the relative size of each of the 29 establishment categories with respect to the others. In addition, by ranking these establishment categories in accordance with their respective median values we can classify them into small and large enterprise sectors. In particular, we classify 9 establishment categories with the smallest median values of establishment size across countries as small enterprise sectors, while the 9 establishment categories with the largest median values are classified as large enterprise sectors. The remaining categories are excluded from our analysis since they may represent the medium enterprise sector, which is not considered in this study. Table 3.1 provides the ranking of manufacturing categories as well as their respective median values of the average number of employees per establishment.

The second stage is the calculation of the relative size of the small enterprise sector. After we reclassify UNIDO's enterprise categories into small and large enterprise sectors, we can calculate their respective size for each country using employment numbers in each sector. In particular, the relative size of the small enterprise sector is calculated as the number of employees in small enterprise sectors as a fraction of the sum of those in small enterprise sectors and those in large enterprise sectors. There are 61 countries where the necessary data are available on an annual basis and cover the 1970-1996 period.

The advantage of this new measure is that it covers a longer period of time than the Ayyagari, Beck and Demirguc-Kunt dataset– it covers the period of 27 years between 1970 and 1996. Furthermore, compared to the previously described official database on SMEs,

---

[16] The year 1985 is selected mainly due to the availability and reliability of the data in that particular year. To check its validity, we also use the year 1990 values as a base for ranking or classifying various establishment categories. There is a strong correlation between the ranking using 1985 as base year and that using 1990 as the base year.

[17] We also calculate the mean values of the average number of employees for each manufacturing categories in order to check the validity of the mean values. The correlation between the mean and the median values is very high.

our new measure has broader coverage of countries. Another advantage of the new measure is that, since only the average number of employees per establishment is taken into account for our classification purpose, it helps reduce the problem associated with the use of any official or unofficial definition of the size of the manufacturing enterprises that often varies from one country to another and from one organization to another.

Another difference between the UNIDO-based and the Ayyagari et al. measures is that whereas the latter considers small and medium enterprises, the first measure takes into account only small and large enterprises and exclude those in between. Therefore, due to these differences, the two measures may differ from one another. In fact, the correlation between them is low at 0.02.

3.3.2 Independent Variables

We consider a wide range of economic, financial, institutional and geographical factors as candidate explanatory variables. Table 3.2 gives a description of all variables used in this study and their respective sources.

The first set of variables are income variables. The log of income per capita (LGDP) is used as a proxy for the level of economic development as well as for market size. We also include its quadratic form ($LGDP^2$) in order to account for a possible non-linear relationship between income and the SME sector share. Moreover, average years of schooling for the population age 15 and over (SCH) is used as a measure of the stock of human capital available in the economy.

Beck et al. (2000) use three indicative measures of financial development: the ratio of commercial bank credit to commercial bank credit plus central bank domestic assets (BANK); the liquid liabilities of the financial system as a percentage of gross domestic product (LLY); and the amount of private credit extended by deposit money banks and other financial institutions as a ratio to gross domestic product (PRIVATE). We construct a new measure of financial development (FINDEV) as the first principal component of these three financial measures.

As for the measure of the degree of financial liberalization we use the Chinn and Ito (2002) index measure of capital account openness (KAOPEN), which is constructed based

on data from the IMF's Annual Report on Exchange Arrangements and Exchange Restrictions.[18] Further discussion of different measures of financial development and financial liberalization will be provided in chapter 4.

In addition, we include a measure of foreign direct investment as a percentage of GDP (FDI) as an indicator of the amount of foreign capital flows into the country. Because foreign investment generally involves a transfer of techniques or product types that imply a relatively large scale of production, we should expect a positive relationship between foreign direct investment and the average firm size.

Kaufmann, Kraay and Zoido-Lobaton (1999) introduce index measures for six different dimensions of governance including voice and accountability, political stability and violence, government effectiveness, regulatory quality, rule of law and control of corruption. In this study, we construct an aggregate measure of institutional efficiency (INST) as an average of these six institutional variables. More discussion of different measures of institutions will be given in chapter 5.

We also use a number of variables representing the regulatory environment that may represent entry barriers to new firms as well as obstacles to growth for existing firms. They include costs of starting up a business as a percentage of GNI per capita (STBUS); log of the costs of obtaining a legal status for a firm to operate as a percentage of GDP per capita (LETRCOSTPC); minimum capital required to start a business as a percentage of GNI per capita (MINCAP); employment laws index of the protection of labour and employment laws (EMPLAW); social security laws index, measuring social security benefits (SSBENF); civil rights index which assesses the degree of protection of vulnerable groups against employment discrimination (CIVILR); and a collective relations laws index which measures of the protection of collective disputes and union power (COLLR).

To account for the potential role of a country's exposure to international markets in determining the distribution of firm sizes we include the log of exports as a percentage of GDP (LEXPGDP) as a proxy for a country's trade with the outside world. We also use

[18] KAOPEN is the first standardized principal component of the IMF's binary variables k_1 (the existence of multiple exchange rates), k_2 (restrictions on current account), k_3 (restriction on capital account transactions) and k_4 (the requirement of the surrender of export proceeds). In order to examine the effect of financial openness – rather than controls – they reverse the values of the AREAER binary variables such that the variable takes a value of zero when a restriction is in place and one when there is no restriction.

other variables that might impair the ability of a country and its firms to trade with the outside world. One of these variables is the average tariff on imported goods (TARIFF) representing a policy distortion that may inhibit cross-border flows of trade. Other variables include the percentage of land area within 100 km of an ice-free coast or navigable river (LND100CR); the ratio of population within 100 km of an ice-free coast or navigable river to total population (POP100CR); the closest air distance to a major international port (AIRDIST) and the dummy for landlocked countries (LANDLOCK), which takes the value of 1 for a country that does not border an open sea. These variables are countries' geographical characteristics which indicate the degree of ease as well as the cost of conducting international trade. For example, a country with a large value for AIRDIST implies that it is situated in a relatively remote location, which could make it much harder for firms in this country to conduct trade with the outside world due to factors such as high transportation costs. This cost issue may, in turn, limit these firms' abilities to grow. Therefore, we would expect a negative relationship between AIRDIST and average firm size.

To determine the potential relationship between infrastructure and the firm size distribution, we follow Herrera and Lora (2005) in constructing an aggregate measure of the quality of infrastructure (INFRSTR) as the first principal component of five different measures. They are: electric power transmission and distribution losses as a percentage of electricity production; percentage of roads that are paved; number of phone lines per 10,000 inhabitants; internet hosts per 1,000 inhabitant; and number of cars per 10,000 inhabitants.

A set of macroeconomic variables is also employed in this analysis. First, the share of the service sector in GDP (SERVGDP) is used as a proxy for the degree of de-industrialization that takes place within an economy. Other variables include the percentage of the urban population in total population (URBPOP) as a proxy for the degree of urbanization, the population growth rate (POPGR), the log of relative labour productivity (LRLP) as a measure of the degree of dualism within an economy, and the annual rate of inflation (INFL) as an indicator of macroeconomic stability.

Additionally, continent dummies for Europe (EU), Sub-Sharan Africa (SAFRI), Latin America (LATAM) and East Aisa (ESEASIA) are used. Finally, we include a dummy variable for oil exporting countries (OIL) to determine whether a country's reliance on oil exports has any effect on firm size structure within the country.

The descriptive statistics of all variables used in this analysis are presented in Table 3.3, while Table 3.4 displays simple correlations between them.

3.4 Framework and Methodology

Bayesian Model Averaging has recently attracted the attention of researchers who are interested in using data to construct economic models to learn about a phenomenon. It was introduced as a method that would enable researchers to avoid a few common issues faced by conventional empirical methods, especially issues associated with model uncertainty.

One of the main problems associated with the conventional approach of hypothesis testing is related to the fact that there are usually many possible candidates for explanatory variables. Researchers normally include many explanatory variables in the regression other than the independent variable of interest itself. The inclusion of such additional conditioning variables aims at minimizing the possibility of the observed associations between the dependent and independent variables being due to other variables and thus misleading.

The choice of which conditioning variables should be included in the regression is normally guided by economic theories. However, a problem arises when the theory is weak and often theoretical arguments point to a rather large number of possible conditioning variables. Such circumstances make it difficult to specify a single model which would best describe the relationships present in the data and thus gives rise to problems of model uncertainty.

The problem of model uncertainty nevertheless has been largely ignored by researchers using conventional empirical methods. Statisticians have traditionally employed a few common approaches that normally involve fine-tuning the model by using various hypothesis testing methods such as t-statistics to select certain variables which are supported by these tests. These variables are then included in the final model, while other variables are dropped. Through these methods, researchers select one model out of many possibilities and proceed as if that was the only model that had ever been considered. It is also common for researchers to attempt to justify their choice by reporting their preferred model accompanied by the results of diagnostic tests – such as tests for outliers, nonlinearity, het-

eroscedasticity or autocorrelation in the residuals – which usually are used to indicate that the chosen model is unlikely to be flawed in a variety of respects.

This approach, however, can yield very misleading results. As Freedman (1983) suggested, by choosing among a large number of models one can increase the probability of finding significant variables by chance alone.[19] Another weakness of the conventional approach, as Granger et al. (1995) points out, is that researchers can arrive at different final models even if they start from the same set of data, simply because of the order in which tests are carried out and the significance levels used. Moreover, the conventional hypothesis testing approach is often criticized because of its reliance on significance tests which typically use a significance level that is arbitrary. Also, though a standard test allows us to say only that the data have failed to reject our null hypothesis of interest, it gives no direct indication of whether the data support the null. A standard test can fail to reject a null hypothesis either because there is not enough data, or because the data support it, but it does not allow us to distinguish between these two different situations (Raftery, 1995).

For these reasons, it is difficult to select with certainty any single model that is superior to all other possibilities. Thus it would be better to present some information about how sensitive the findings are to alternative modeling choices.

In this sense, Bayesian Model Averaging methods based on the ideas proposed by Leamer (1978) provide a good solution to this problem, by computing the quantities of interest such as parameter estimates through averaging across a large set of possible models. Until recently, however, the Bayesian ideas were rarely used in practice due to difficulties associated with their implementation. Recent improvements in computing power as well as further theoretical developments have significantly facilitated the use of these ideas and thus increased their popularity among researchers.

The basic idea behind Bayesian Model Averaging is to average over a number of alternative models instead of presenting just one particular model and assuming it to be the true model. It may provide better predictive ability than the conventional approach. The Bayesian approach treats all unknown parameters as random variables and expresses

[19] On a similar note, Leamer (1983) shows that standard model selection methods can lead to a regression with high R^2 and a highly significant F-statistic and thus lead researchers to believe that there is some sort of relationship between dependent and various explanatory variables present in the data even when in reality there is none.

all uncertainty in terms of a probability distribution. Likewise, it treats the true model as unknown, and then attempts to summarize the uncertainty about the model in terms of a probability distribution over the model space. Priors are assigned to the models as well as to the parameters and posterior probabilities for different models are calculated based on the data. Posterior distributions for the parameters can be calculated by averaging the posterior distributions under each model weighted by their corresponding posterior model probabilities.[20]

To illustrate this, let us look at the case with N possible models which we denote as M_1, ... , M_N, with the corresponding parameter vectors $\theta = (\theta_1, \ldots, \theta_N)$. The Bayesian approach to model uncertainty is to assign a prior probability to each model, $p(M_n)$, and a prior probability distribution $p(\theta_n \mid M_n)$ to the parameters of each model.[21]

If Δ is our parameter of interest, then the full posterior distribution of Δ given the data D, i.e. $p(\Delta \mid D)$, is calculated as the average of the posterior distributions under each model weighted by the corresponding posterior model probabilities. It is given by:

$$p(\Delta \mid D) = \Sigma_{n=1}^{N} p(\Delta \mid M_n,\ D)\ p(M_n \mid D) \tag{3.1}$$

where $p(\Delta \mid M_n,\, D)$ is the posterior distribution obtained under a given model and $p(M_n \mid D)$ is the posterior model probability, which is the posterior probability that M_n is the correct model after having observed the data D, and given a certain prior over models.

This approach requires us to calculate the posterior model probability, which is given by:

$$p(M_n \mid D) = \frac{p(D \mid M_n)\ p(M_n)}{\Sigma_{k=1}^{N} p(D \mid M_k)\ p(M_k)} \tag{3.2}$$

where

$$p(D \mid M_n) = \int p(D \mid \theta_n,\ M_n)\ p(\theta_n \mid M_n)\ d\theta_n \tag{3.3}$$

[20] More discussion on this subject can be found in Raftery (1995); Hoeting, Madigan and Raftery (1997); Hoeting et al. (1999); Temple (2000); and Fernandez et al. (2001).

[21] This discussion draws heavily on Raftery (1995).

is the marginal (or integrated) likelihood for model M_n, θ_n is the vector of parameters of model M_n, $p(\theta_n \mid M_n)$ is the prior probability distribution of θ_n under model M_n, $p(D \mid \theta_n, M_n)$ is the likelihood of θ_n under model M_n, and $p(M_n)$ is the prior probability that M_n is the true model (Ratery, Madigan and Hoeting, 1997).

To determine the preferred models, we need to evaluate posterior model probabilities. This requires us to calculate the *Bayes factor*, a quantity which enables us to decide which model is better supported by the data.[22]

To illustrate the Bayes Factor, we assume that there are only two competing models M_1 and M_2 with parameter vectors θ_1 and θ_2. In this case, the posterior probability in equation (3.2) can be rewritten as:

$$p(M_1 \mid D) = \frac{p(D \mid M_1)\, p(M_1)}{p(D \mid M_1)\, p(M_1) \;+\; p(D \mid M_2)\, p(M_2)} \tag{3.4}$$

where $p(M_1 \mid D)$ is the posterior probability that M_1 is the correct model. The same expression holds for $p(M_2 \mid D)$, the posterior probability that M_2 is the correct model. By construction, $p(M_1 \mid D) + p(M_2 \mid D) = 1$.

The extent to which the data support one model over the other, say M_2 over M, is given by the *posterior odds* for M_2 against M_1. The posterior odds is effectively the ratio of the posterior probabilities of these models, which can be written as:

$$\frac{p(M_2 \mid D)}{p(M_1 \mid D)} = [\frac{p(D \mid M_2)}{p(D \mid M_1)}]\, [\frac{p(M_2)}{p(M_1)}] \tag{3.5}$$

The first factor on the right-hand side is the Bayes factor for M_2 against M_1, denoted as B_{21}. The second factor on the right-hand side is the prior odds. Without any prior preference, each model is assumed to have the same probability of being the true model, that is $p(M_1) = p(M_{2)} = \frac{1}{2}$. Thus, equation (3.5) becomes:

[22] See Kass and Raftery (1995) for more discussion of the Bayes factor.

$$\frac{p(M_2 \mid D)}{p(M_1 \mid D)} = \frac{p(D \mid M_2)}{p(D \mid M_1)} \tag{3.6}$$

which implies that the posterior odds depend only on the Bayes factor. If $B_{21} > 1$, then the data support M_2 over M_1.

Unfortunately, evaluating the Bayes factor involves calculating the integrated likelihood as given in equation (3.3), which can be difficult especially when there are many possible models. Over the years, various analytical and numerical approximation methods have been proposed to address this problem. Schwarz (1978) proposes the use of a Bayesian Information Criterion (BIC) to approximate the Bayes factors that are needed to compute the posterior model probabilities.[23]

For linear regression with normal errors, the BIC takes the following form:

$$BIC'_k = n \ \log(1 - R^2_k) + p_k \ \log \ n \tag{3.7}$$

where R^2_k is the R-squared for model M_k and p_k is the number of independent variables, excluding the intercept. BIC'_k assesses how well model M_k can predict the data, given a number of explanatory variables. According to the BIC approximation, a model with a higher R^2 and fewer parameters – which results in a lower BIC'_k value – is considered a preferred model.

If we assume that all N models are equally likely before examining the data, so that $p(M_k) = \frac{1}{N}$ for all k, then using the BIC approximation to twice the logarithm of the Bayes factor, equation (3.2) can be rewritten as:[24]

$$p(M_n \mid D) \approx \frac{\exp(-\frac{1}{2} \ BIC'_n)}{\Sigma_{k=1}^{N} \exp(-\frac{1}{2} \ BIC'_k)} \tag{3.8}$$

Estimating every single model is not possible due to the fact that there are too many possible models to consider. In this case, there are as many as 31 candidate explanatory variables and, thus, 2^{31} (or more than two billion) possible models to estimate.

[23] See also Raftery (1995), for example, for a more detailed discussion on the technical aspects of the derivation of BIC approximation as well as the estimation and interpretation of the Bayes Factor.

[24] See Malik and Temple (2004). Also see Raftery (1995) and Sala-i-Martin et al. (2004).

A possible solution to this problem is the use of the Occam's Window approach as proposed by Madigan and Raftery (1994). One version of Occam's Window involves excluding from the averaging procedure any model that is much less likely than the model with the highest posterior model probability. This is called the *symmetric* version of Occam's Window, which is used in this analysis. For the purpose of this study, we only select models with the posterior model probability of at least 0.20. A stricter version of Occam's Window does not only exclude models that fall under the rejected criteria of the symmetric version, but also drops all sub-models that are nested within these excluded models.[25] The *BICREG* software for the R statistical language is used to implement the Occam's Window procedure.[26]

The Markov chain Monte Carlo model composition (MC^3) approach of Madigan and York (1995) is also experimented with, in addition to the Occam's Window approach.[27] This approach uses a Markov chain Monte Carlo method to approximate all models in equation (3.1). Generally, the results from the MC^3 approach are more outlier-robust than those from the Occam's Window approach. To implement this procedure, the *MC3.REG* software in R is used.[28]

In our study, Bayesian Model Averaging is used to identify the models that have high explanatory power as reflected in their high posterior model probabilities, as well as to identify variables that have high posterior probabilities of being included. Therefore these two quantities are vital in our model selection process. As mentioned earlier, the posterior model probability is used to compare different models by assessing the degree of support each model receives from the data. In other words, it tells us the probability of a particular model being the true model after seeing the data. Similarly, the posterior probability of inclusion for a certain variable is used to identify which variables are worthy

---

[25] Madigan and Raftery (1994) provide a detailed description of the algorithm and show how averaging over the selected models provides better predictive performance than basing inference on a single model in each of the examples they consider.

[26] The software was initially written for the S-Plus language by Adrian Raftery and revised by Chris Volinsky. It was later modified for the R language by Ian Painter.

[27] See Hoeting et al. (1997) for more discussion of this approach.

[28] Like the *BICREG* software, the *MC3.REG* software was developed for the S-Plus language and later modified for the R language.

of being included in the models. It is defined as the sum of the posterior probabilities of all the models in which this variable appears, i.e. where its coefficient is non-zero.

In addition to reporting the results from the averaging exercises, we present Ordinary Least Square estimates of some of our preferred models, in order to determine the robustness of these models by subjecting them to various conventional diagnostic tests. This method will also allow us to investigate the magnitude of the effects of various explanatory variables on the dependent variable.

3.5 Empirical Results

In this section we present the results from the Bayesian Model Averaging exercises as well as from Ordinary Least Square regression estimation of the top models suggested by the BMA.[29]

The results are divided into two sections using the share of the small and medium enterprise sector, SMEOFF, and the UNIDO-based measure of the relative size of the small enterprise sector versus the large enterprise sector, SEMSELE, as the dependent variable, respectively.

We use the model averaging method to determine the variables to be included in the models by calculating their respective posterior probabilities of inclusion. We also provide a sign certainty index, which is evaluated based on the sum of posterior model probabilities for all models in which a variable acts in a given direction (e.g. negative). The sign certainty index enables us to see the likely direction of the relationship between the dependent variable and the explanatory variable under consideration. In addition, we present the results from an outlier-robust BMA in order to check the robustness of these relationships.[30]

---

[29] We conduct the Bayesian Model Averaging exercises using the R statistical program. The *BICREG* command function is used to perform this task.

[30] The outlier-robust BMA can be performed using the *MC3* function in R.

3.5.1 BMA Results for SMEOFF

The analysis is conducted using data from 76 developing and developed countries and a set of 31 candidate explanatory variables. Table 3.5 provides a list of the sample countries. The dependent variable is the share of the small and medium enterprises in manufacturing sector, SMEOFF. All variables are averaged over the 1990-99 period unless noted otherwise. We impute missing data for a small number of variables using a regression-based imputation method. The results are presented in Table 3.6.

We experiment with different set of explanatory variables. Column 1 includes indicators of income, human capital and a number of macroeconomic variables as explanatory variables. In column 2 we add the log of the ratio of exports to GDP as well as other trade-related variables The measures of financial development and financial openness are added in column 3, while in column 4 we include various indicators of institutions. In column 5 we conduct a BMA exercise with a full set of candidate explanatory variables as in column 4, but we drop both Hong Kong and Singapore from our sample due to the possibility of them being outliers.[31] Finally, in column 6 we present the posterior inclusion probability for each explanatory variable obtained by using outlier-robust BMA. In all cases, any variable with a posterior inclusion probability (PIP) of 20 percent or greater is considered important.

The first relationship to look at is that between firm size distribution and income level. Figure 3.1 shows the simple relationship between the SME sector share and income. The results from the BMA exercise in columns 1-5 indicate that both the log of GDP per capita, LGDP, and its quadratic form, $LGDP^2$, have consistently high posterior probabilities of being included in the models. Results from the outlier-robust BMA in column 6 also confirm the importance of these variables. Furthermore, the coefficients on both terms have opposite signs as reflected by their sign certainty indexes, with the log of income and its squared value having positive and negative signs, respectively.[32] These signs suggest there exists a non-linear, inverted-U shape relationship between the SME sector share and the level of income. This finding is consistent with the hypothesis that there is a decline in the

[31] This is apparent in the case of the ratio of exports to GDP as shown in Figure 3.

[32] The only exception is the the sign certainty index for the squared value of income in *column 1* where it comes in as uncertain although its posterior inclusion probability is still fairly high.

relative importance of the small and medium enterprise sector in the economy over the long run. In the short run, however, the results indicate a relative increase in the share of small and medium enterprises in the economy. This phenomenon may be due to an increase in entry by small new firms as well as firms moving from small to medium size, perhaps in response to certain new opportunities that take place at the earlier stage of development. These developments would expand the size of the small and medium enterprise sector.

Nevertheless, at a later stage of development when new opportunities become scarce and competition is fierce, only a small number of firms are able to grow and expand while others remain small or even disappear. This leads to an increase in concentration in a small number of large firms. Moreover, other activities such as takeovers and merger and acquisitions that often take place in developed economies may also shift the firm size distribution towards larger firms.

Unlike some previous studies, we do not find any significant relationship between our measure of human capital, SCH, and the relative importance of the SME sector. The posterior inclusion probability for SCH is low in almost all cases, except in column 5 when potential outliers are taken into account. In this case, results suggest a negative relationship between the human capital measure and the share of the small and medium enterprise sector. Figure 3.2 shows the scatter plot of this relationship.

Our trade variable, LEXPGDP, is found to have high explanatory power as reflected by its high inclusion probability. Furthermore, the negative sign certainty index suggests an inverse relationship with the SME sector share, which implies that a country which has more exposure to international trade tends to have large firms. A plausible reason for this negative relationship is that firms in countries with an open trading system may find it easier to obtain access to larger global markets that is positive for their growth and expansion. This is consistent with the findings of studies by Wagner (1995, 2001), which show a positive relationship between firm size and direct export activities.

Figure 3.3 shows a simple scatter plot of the relationship between LEXPGDP and the dependent variable. A brief look at this figure seems to suggest a potential outlier problem, and thus we must treat our BMA results with caution. To address this issue, in column 5 we drop both Singapore and Hong Kong from our sample countries due to their potential role as outliers. However, dropping these countries does not have much effect on the results. The

relationship between LEXPGDP and the dependent variable is still negative and significant. In addition, results from outlier-robust BMA, which takes into account potential bias due to outliers, in column 6 also shows this relationship to be robust, with a high probability of inclusion of about 83 percent.

Additionally, columns 2 and 3 show that AIRDIST, a measure of a country's geographical proximity to major international markets, has a significant relationship with the dependent variable. Its positive sign certainty index suggests that a country that is located in a more isolated location tends to have a larger share of small and medium enterprises in the economy. This might result from the difficulties faced by firms in this country in conducting trade with the outside world due to such factors as high transportation costs, and thus limit their abilities to exploit economies of scale. Nevertheless, this relationship is not robust to the inclusion of institutional indicators in column 4. Outlier-robust BMA analysis also shows the fragility of this relationship.

As for financial factors, the results from BMA exercises do not show any significant role of either financial development or financial openness in explaining cross-country variation in the firm size distribution, as proxied by the employment share of small and medium enterprises in the manufacturing sector. Similarly, the share of foreign direct investment in GDP, FDI, which is a proxy for the amount of capital flows into the country does not have any significant effect either.

We now examine the role of institutions. Figure 3.4 displays the relationship between our aggregate index measure of the quality of institutions, INST, and the SME sector share. It shows a positive relationship between these two variables. Results from BMA in column 4 confirm this positive relationship. Moreover, high posterior inclusion probabilities for our institutional quality measure suggest its important role in explaining cross-country differences in the share of the SME sector. In addition, outlier-robust BMA results in column 6 show the robustness of this relationship. These findings imply that a country with more effective and efficient institutions tends to have a business environment that is more conducive to the growth of small and medium enterprises. As explained earlier, good institutions may help level the playing field for firms of all sizes and thus gives small and medium sized firms a better opportunity to compete with larger firms and to grow.

On the other hand, with the exception of the employment laws index, EMPLAW, all other indicators of regulatory environment do not come in as relevant. BMA results in column 4 show a negative relationship between the employment laws index and the share of the SME sector, which suggests that strict employment laws have stronger adverse effects on firms of small and medium sizes relative to larger ones. However, this relationship is not robust when outlier-robust BMA is applied. Its inclusion probability drops to 3 percent in column 6.

The effect of a nation's physical infrastructure, INFRSTR, on the firm size distribution is ambiguous. The results from a BMA analysis in columns 1-4 do not show any significant relationship between INFRSTR and the share of the SME sector. However, the relationship turns out to be significant when the outlier-robust BMA analysis is conducted, with an inclusion probability of 43 percent. Hence, further investigation should be conducted in order to determine the robustness as well as the direction of this relationship.

Regional dummies for Europe, Latin America and East Asia have important explanatory powers as reflected by their high inclusion probabilities. The positive sign certainty indexes on these variables suggest that small and medium enterprises play important roles in economies across various regions of the world.

Finally, the results in columns 1-3 show that macroeconomic stability, as proxied by the annual inflation rate (INFL), has a negative and significant relationship with the dependent variable. This may suggest that macroeconomic instability has a disproportionately large, adverse effect on small and medium enterprises compared to their larger counterparts. One potential explanation is that large firms may have a much better ability than smaller ones in generating the internal capital needed to withstand economic shocks and thus have a better chance of survival during hard times compared to their smaller counterparts. This finding is in contrast with that of Mulhern and Steward (2003) who suggest that the share of small and medium enterprises tends to increase during recessions because large firms suffer more during hard times.

It is, however, important to note that the explanatory power of INFL decreases with the addition of indicators of institutions in column 4. The relationship is also found to be non-robust when the outlier-robust BMA method is applied.

Table 3.7 displays the structures of the top ten models, ranked in accordance with their respective posterior model probability (PMP), from the BMA exercise in column 4 when all explanatory variables are included. The best model contains seven explanatory variables, with a PMP of almost 0.08. This is much higher than the prior model probability, considering there are 2^{31} potential models to choose from.

3.5.2 OLS Results for SMEOFF

In order to further examine the robustness as well as the magnitude of the relationship between different explanatory variables and our indicator of the firm size distribution, we present ordinary least squares regression results for the top 10 models suggested by the BMA exercise.[33] Results are presented in Table 3.8.

The results confirm the relationship between income level and the relative importance of small and medium enterprises in the economy. The coefficients of income and its square are statistically significant at 1 percent level in all the cases. In addition, we find that the signs of their coefficients are as expected. We can estimate the turning points – that is the income levels where we may expect the SME share to begin to decline. Based on these top ten models, we estimate that the share of small and medium enterprises in the economy is likely to begin to reverse its course from rising to declining when income per capita reaches a level roughly between 1,900 and 3,000 US dollars. These estimations of the turning points as well as the number of countries that lie to the left of these turning points are also presented in the table.

Results in column 5 point to a possible negative relationship between our measure of human capital and the dependent variable, with the relationship significant at the 5 percent level. However, this relationship is rather fragile. For instance, the inclusion of employment protection laws in column 10 causes this relationship to become insignificant. Therefore, as in the BMA exercises, we do not find any robust relationship between human capital and the relative importance of the SME sector.

[33] It is important to note that these top ten models are selected based on the BMA exercise in column 4 of Table 3.6 when all explanatory variables are included. In addition, a number of variables such as INFL, INFRSTR and various financial indicators will not be discussed here as they are not included in these top ten models.

Table 3.8 also shows a strong explanatory role for the continental dummies. We find that all three dummy variables for Europe, Latin America and East Asia have statistically significant and positive relationships with the dependent variable.

Ordinary least square results confirm the findings from BMA which suggest that there exists a significantly negative relationship between a country's exposure to international markets, as proxied by its share of exports in GDP (LEXPGDP), and the relative importance of the SME sector. The t-statistics indicate that this relationship is significant at the 1 percent level in all cases. The size of the coefficients on LEXPGDP further suggest not only a statistically significant but also an economically meaningful relationship. For example, using the model in column 1, it is estimated that a 1 standard deviation increase (decrease) in the log of the share of exports in GDP would result in a decrease (increase) of 0.31 of a standard deviation in the relative importance of the SME sector in the economy.

Additionally, results in column 6 indicate a possible positive relationship between AIRDIST and the share of the SME sector. Such a positive relationship may suggest that countries situated in relatively remote locations tend to have predominantly small firms in their economies. As explained earlier, one potential explanation for this phenomenon is that such geographical isolation creates barriers for firms in these countries in accessing markets beyond their own borders, which in turn limits their abilities to expand and prevents them from exploiting potential economies of scale.

With regards to the role of institutions, OLS results show that the quality of institutions has a positive relationship with the relative importance of the SME sector. The t-statistics and the size of the coefficients on INST show that this relationship is both statistically and economically significant. For instance, a 1 standard deviation change in the quality of institutions is associated with a change of 0.62 of a standard deviation in the relative share of the SME sector.

Moreover, results in Table 3.8 indicate a negative relationship between a measure of the regulatory environment – the employment protection laws, EMPLAW – and the dependent variable, which is consistent with the findings from the BMA exercises.

Finally, whereas our BMA results do not show the rate of population growth, POPGR, to have a robust relationship with the SME sector share, OLS results indicate this relationship to be positive and significant. Nevertheless, as explained earlier, because of various

issues commonly associated with simple cross-section analysis such as the OLS method, we should be careful in interpreting such a result. Hence, further investigation is needed before any concrete conclusion can be drawn about this relationship.

Overall, the findings from ordinary least squares regression estimates of the top ten models are consistent with the BMA results regarding the relationship between a number of explanatory variables and the share of the SME sector in the economy.

3.5.3 BMA Results for SEMSELE

In this section, we conduct a BMA analysis using the UNIDO-based measure of the relative importance of the small enterprise sector versus the large enterprise sector, SEMSELE, as the dependent variable. This is done primarily to determine the relationship between various candidate variables and enterprises of different sizes, namely small and large enterprises.

The data are from 53 developing and developed countries and a set of 30 candidate explanatory variables.[34] All variables are averaged over the 1990-96 period unless indicated otherwise. A list of the sample countries used in this case is presented in Table 3.9. We present results in Table 3.10. As explained earlier, any variable with a posterior inclusion probability (PIP) of 20 percent or higher is considered important.

Column 1 includes indicators of income, human capital and a number of macroeconomic variables as explanatory variables. Initial results suggest that neither our measure of income (LGDP) nor its squared value ($LGDP^2$) has any significant explanatory power for cross-country differences in the relative size of the small enterprise sector versus the large enterprise sector. Similarly, the relationship between the measure of human capital (SCH) and the dependent variable is shown to be insignificant. Figure 3.5 displays the relationship between human capital and the relative size of the small enterprise sector.

Among variables that have high probabilities of inclusion are the dummies for countries in Africa and Latin America. Also, the dummy for oil exporting countries seems to

[34] Countries such as Hong Kong and Singapore which may be potential outliers are not included in the sample countries.

have a positive and significant relationship with the dependent variable as shown by its high inclusion probability and its positive sign certainty index.

In addition, results suggest that the share of the service sector in GDP (SERVGDP), a proxy for the degree of de-industrialization that takes place within an economy, has a positive and significant relationship with the dependent variable. This finding is as expected because in general service sector firms tend to be of limited scale and specialized in nature due to the types of demand for their products and services. A graphical representation of the relationship between the service sector share in GDP and the relative size of the small enterprise sector versus the large enterprise sector is shown in Figure 3.6.

In columns 2 and 3 we add trade-related variables and financial variables, respectively, to our list of explanatory variables. The explanatory power of the human capital measure increases with the addition of these new variables. The posterior probability of inclusion for human capital increases from around 5 percent to over 39 percent. Further, sign certainty indexes indicate a negative relationship, which suggests that a country with a high level of human capital tends to have its economy dominated by larger firms. This is consistent with the view that a greater availability of qualified human resources makes it possible for firms to specialize in more complex goods and acquire technologies that demand larger and more complicated production processes.

The inclusion of these new candidate variables also increases the explanatory power of the share of urban population in total population (URBPOP) and its PIP increases from just under 10 percent to 27 percent. Sign certainty indexes for this variable suggest a positive relationship with the dependent variable, implying that a higher degree of urbanization is associated with a larger role of the small enterprise sector in the economy. Figure 3.7 displays the graphical relationship between these two variables.

The log of exports over GDP (LEXPGDP) also has a great deal of explanatory power for the cross-county differences in the relative size of the small enterprise sector versus the large enterprise sector, as indicated by its consistently high probability of inclusion. Sign certainty indexes suggest a positive relationship between them. A scatter plot of this relationship is shown in Figure 3.8. This finding, however, is inconsistent with earlier findings that indicate a negative relationship between the ratio of exports to GDP and the share of the small and medium enterprise sector. A possible explanation for this inconsistency rests

on the differences between the two measures of small and medium enterprises – SMEOFF and SEMSELE – in terms of their construction, as explained earlier.

Another potential reason for the inconsistency in the findings concerns the extent of interfirm cooperation within a country. As mentioned in the previous chapter, in countries such as Korea, there exists dynamic interfirm cooperation between firms of different sizes through such arrangements as sub-contracting networks. Under such an arrangement, large firms rely heavily on small firms for supplies of intermediate products. Therefore, in these countries, expansion in international trading activities is likely to increase the role of small firms instead of reducing it. This is consistent with the finding when the relative size of the small enterprise sector, SEMSELE, is used as the dependent variable. Nevertheless, due to insufficient information on the extent of interfirm cooperation in each individual country, we do not know how many countries there are in our sample in which such interfirm co-operation plays significant roles. Thus, it is impossible to assess to what extent our results may be affected by such a factor.

Results also show a positive and significant relationship between a country's closest air distance to a major international port (AIRDIST) and the dependent variable. Figure 3.9 presents a simple graphical view of this relationship.

Additionally, among our dummy variables, only the relationship between the dummy for countries in Latin America (LATAM) and the dependent variable stays robust. The explanatory powers of the dummies for countries in Africa (SAFRI) and for oil exporting countries (OIL) decrease with the inclusion of these new conditioning variables.

The inclusion of institutional variables in column 4 does not change our earlier findings. The probabilities of inclusion for human capital, the share of urban population, the ratio of exports to GDP and for air distance still suggest that they have a significant effect on the dependent variable. Moreover, the sign certainty indexes on these variables remain unchanged.

As for our institutional measures, only the indicator of employee rights (CIVILR) has a significant relationship with the relative size of the small enterprise sector versus the large enterprise sector. Furthermore, the sign certainty index indicates a negative relationship between CIVILR and the relative size of small enterprises. Moreover, none of our measures of financial factors seem to have a significant relationship with the dependent variable.

Finally, in column 5 we apply the outlier-robust BMA method to check the robustness of earlier results. Due to software limits on the number of explanatory variables that can be included at any one time, some of the explanatory variables used in column 4 are excluded from the outlier-robust exercise in column 5.[35] Specifically, the excluded variables consist of the dummy variable for oil exporting countries (OIL), the percentage of land area within 100 km of an ice-free coast or navigable river (LND100CR), the ratio of population within 100 km of an ice-free coast or navigable river to total population (POP100CR), the social security laws index (SSBENF), and the collective relations laws index (COLLR). They are selected because they consistently have very little explanatory power as shown in columns 1-4.

The application of the outlier-robust method does not have much effect on our findings, which suggest that our results are not excessively influenced by outlying observations. Of the seven variables that previously had high explanatory power, only the share of urban population in total population (URBPOP) becomes insignificant: the inclusion probability drops to 6 percent.

The structures of the top ten models from the BMA exercise are presented in Table 3.11.[36]

3.5.4 OLS Results for SEMSELE

Table 3.12 presents ordinary least squares regression results of the above mentioned top ten models. The relationship between the human capital measure (SCH) and our dependent variable, the relative size of the small enterprise sector versus the large enterprise sector (SMESELE), is shown to be negative and significant. This relationship is robust to changes in the set of other conditioning variables as shown in columns 4, 6 and 8. Furthermore, the sizes of the coefficients on human capital indicate a fairly strong correlation between human capital and the dependent variable.

---

[35] The statistical program for R which is used to conduct this exercise requires that the number of explanatory variables used does not exceed half of the number of observations.

[36] The models presented are based on the BMA exercise in column 4, when a full set of conditioning variables is used.

Results in columns 4 and 6 show that the relationship between the degree of urbanization as proxied by the urban population share of the total population (URBPOP) and the dependent variable is positive and statistically significant at the 5 percent confidence level. Nevertheless, the size of the coefficients on this variable show the magnitude of this relationship to be rather small. Similar conclusions can be drawn about the relationship between the share of the service sector in GDP (SERVGDP) and air distance to a major international port (AIRDIST) on the one hand and our dependent variable, the relative size of the small enterprise sector versus the large enterprise sector, on the other.

Among our candidate variables, the log of export share in GDP (LEXPGDP) has statistically the most robust relationship with the dependent variable with it being the only variable included in all the top ten models. Furthermore, the size of coefficients on this variable are fairly large suggesting a strong relationship with the share of the small enterprise sector.

Finally, columns 5 and 7 point out that the relationship between the dependent variable and our index measure of employees' rights (CIVILR) is negative and significant at 10 percent level. The magnitude of this relationship is fairly large judging from the sizes of the coefficients. However, the significance of this variable decreases with changes in the set of other conditioning variables as shown in column 10, indicating that the relationship is fragile.

3.6 Conclusion

This chapter uses a fresh empirical method to investigate various possible determinants of the firm size distribution. Bayesian model averaging procedures are applied using a large number of candidate explanatory variables. Two separate dependent variables are employed, namely the Ayyagari et al. (2003) measure of the share of the small and medium enterprise sector in the economy and the UNIDO-based measure of the relative size of the small enterprise sector versus the large enterprise sector.

Empirical results indicate a non-linear relationship between the level of income and the share of the SME sector in the economy, where the SME sector rises in relative impor-

tance initially and then declines at a later stage of development. It is also found that better and more effective functioning of institutions helps to create an environment in which smaller firms find it easier to compete and to grow. Yet, not all regulations exert such positive effects. For instance, regulations such as employment protection laws are found to have a disproportionately negative effect on small and medium size firms compared to larger firms.

The relationship between the trade variable and the dependent variable is ambiguous. On the one hand, the share of exports in GDP is found to have a negative and significant relationship with the SME sector share. On the other hand, the relationship becomes positive when the UNIDO-based measure is used as the dependent variable. Moreover, unlike previous studies, we do not find any robust relationship between the relative importance of the SME sector, on the one hand, and measures of human capital, finance, and physical infrastructure, on the other.

Finally, it is important to note that the results from the BMA exercise in this paper only show potential associations between variables of interest and the dependent variable. They do not necessarily imply any causal relationship. To investigate potential causality, other empirical methodologies are needed, and this will be undertaken in later chapters.

Appendix 3.1: Sectoral Classification

This section describes the procedure used to classify various establishment categories into small and large enterprise sectors based on a dataset obtained from the United Nations Industrial Organization (UNIDO). The UNIDO dataset is classified into 29 different establishment categories and there are 65 countries where data are available. Our classification procedure involves four simple steps.

Step 1: Calculate average number of employees per establishment

The first step in our sectoral classification process is to calculate average number of employees per establishment. This calculation can be done as follow:

Let $N_{i,j,t}$ and $E_{i,j,t}$ be the total number of establishments and total number of employees for each category, where subscripts i, j and t denote each particular sector, country, and year, respectively.

Using UNIDO data, we calculate the average number of employees per establishment for each country and sector by dividing the total number of employees by the number of establishments for each respective country in that particular year. Thus, for each country:

$$AVG_{i,j,t} = \frac{E_{i,j,t}}{N_{i,j,t}} \quad \text{for } i = 1 \text{ ... } 29,\ j = 1,\ 2,\ \text{... },65 \text{ and } t = 1970, ..., 1996$$

where $AVG_{i,j,t}$ denotes average number of employees per establishment for UNIDO industrial sector i, in country j and in year t.

It is, however, important to acknowledge that while UNIDO industrial dataset provides data for the number of employees in each industrial category, E, from 1970 onwards, data for the number of establishments, N, are

available only from 1981 onwards. Therefore, only values for average number of employees per establishment between 1981 and 1996 can be calculated.

One issue of concerns relates to the missing annual values of average number of employees per establishment. To deal with this issue, we fill in the missing values by using a linear-interpolation method so as to maximize the number of observations for our econometric analysis of the dataset. Linear interpolation is a simple form of interpolation which imputes data points between two existing (known) data points in a linear fashion – analytically speaking using linear polynomials, geometrically speaking on a straight line between two known points. In our study, for a particular country, the linear interpolation method is used to generate missing value(s) of average number of employees per establishment for a sector and in a particular year(s) based on the values of the average number of employees per establishment in that particular sector in other years where data are available.

Table 3.13 displays countries and years where the data for average number of employee per establishment are missing and are thus imputed.

For this study, we use the year 1985 ($t = 1985$) as our base year for sectoral classification. Thus, for each country, average number of employee per establishment in 1985 can be calculated as follow:

$$AVG_{i,j,1985} = \frac{E_{i,j,1985}}{N_{i,j,1985}} \quad \text{for } i = 1 \text{ ... } 29 \text{ and } j = 1, 2, \text{ ... }, 65$$

where $AVG_{i,j,1985}$ is average number of employees per establishment in 1985 for sector i and in country j.

Table 3.14 presents summary statistics for the average number of employees per establishment in 1985. Moreover, we examine data for the years 1981 and 1995, two years close to both ends of our study period and where data are generally available. Table 3.15 and 3.16 show summary statistics

for average numbers of employees per establishment for years 1981 and 1995, respectively.

Additionally, for these three years we present density function plots of the 29 establishment categories so that it can be seen which sectors are large and whether there is significant difference between the size of establishments in different countries. Figures 3.10, 3.11 and 3.12 display density function plots of the 29 industrial sectors for 1981, 1985 and 1995, respectively.

Step 2: Calculate median value

The next step in our sectoral classification is calculating the median values of the average number of employees per establishment for each category. In so doing, we can compare across sectors the average size of industrial sectors based on average number of employees per establishment.

Thus, for each sector:

$$MEDIAN_{i,t} = Median\ [AVG_{i,j,t}] \quad \text{for } i = 1 \ldots 29$$

where $MEDIAN_{i,t}$ is the median value of the average number of employees per establishment for industrial category i in year t.

Figures 3.13 a and b display time series plots for the median values of the average number of employees per establishment for the 29 industrial sectors between 1981 to 1996. Here, time series plots are broken into two separate graphs because of the large variation across sectors in the median values of the average number of employees per establishment.

Again, for the purpose of this study, we chose 1985 as the base year for our calculation of the median values for the average number of employees per establishment. Thus, for each industrial category:

$$MEDIAN_{i,1985} = Median\,(AVG_{i,j=1,1985}, AVG_{i,j=2,1985}, ..., AVG_{i,j=65,1985})$$
for $i = 1$ ... 29

where $MEDIAN_{i,1985}$ is the median value of the average number of employees per establishment for category i in 1985.

Step 3: Sort median value and classify enterprise sectors

Each establishment category is then ranked in accordance with their respective median values of average number of employees. Once this ranking is done, we can classify various establishment categories into the small enterprise sector, i.e. those dominated by small enterprises, and the large enterprise sector, those dominated by large enterprises.

In this case, we classify 9 establishment categories with the smallest median values of establishment size across countries as the small enterprise sector, while the 9 establishment categories with the largest median values are classified as the large enterprise sector. The remaining categories are excluded from our analysis since they may represent the medium enterprise sector, which is not considered in this study. Table 3.1 provides the ranking of manufacturing categories as well as their respective median values of the average number of employee per establishment.

Furthermore, in order to check the relevancy of using the median value of average number of employees per establishment as a basis for our sectoral classification, we produce two other measures of business sectors and use them as basis for new sectoral classifications. The two new measures are mean-of ratios (MOR) and ratio-of-means (ROM). For the year 1985,

$$MOR_{i,1985} = \sum_{j=1}^{65} (E_{i,j,1985}/N_{i,j,1985}) \quad \text{for } i = 1, ..., 29$$

$$ROM_{i,1985} = \sum_{j=1}^{65} E_{i,j,1985} / \sum_{j=1}^{65} N_{i,j,1985} \quad \text{for } j = 1, ..., 29$$

Once these new measures have been calculated, we proceed to classify industrial sectors based on these new values. Table 3.17 and 3.18 present sectoral classifications based on values of mean-of-ratios (MOR) and ratio-of-means (ROM), respectively.

We then compare these new classifications to the classification based on the median value of average number of employees per establishment. A high degree of correlation between these different sectoral classifications would strongly support the use of the median value of average number of employees per establishment as a basis for sectoral classification.

Table 3.19 displays together values of the median of average number of employees per establishment (or median-of-ratio), mean-of-ratios (MOR) and ratio-of-means (ROM). The correlations between median values of ratios, on the one hand, and mean-of-ratios and ratio-of-means, on the other, are fairly high at 0.96 and 0.86, respectively. Such high correlations support the use of median values of average number of employees per establishment as a base for classification of manufacturing sectors. Moreover, Table 3.20 compares sectoral classifications using all three measures of business sectors.

Step 4: Calculate the relative employment share of the small enterprise sector

Once the classification of sectors is done, we can use UNIDO data on sectoral employment to calculate the total employment in the small enterprise sector as well as the large enterprise sector. For example, the total employment in the small enterprise sector equals the combined number of employees in the 9 establishment categories which were classified as the small enterprise sector. The same applies to the calculation of the total employment in the large enterprise sector.

Then, for each country, we can calculate annual values of the relative

employment share of the small enterprise sector versus the large enterprise sector, denoted as SEMSELE, as follows:[1]

$$SEMSELE = \frac{SE\ Employment}{SE\ Employment + LE\ Employment}$$

where *SE Employment* is the total number of employees in the small enterprise sector and *LE Employment* is the total number of employees in the large enterprise sector.

Comparison between SMEOFF and SEMSELE

This new measure has a number of advantages over previous measures of small business sector. One of the most important of these advantages is that the new measure covers a long period of time compared to most previous measures – covering 27 years between 1970 and 1996. Furthermore, compared to the previously described official database on SMEs, our new measure has broader coverage of countries. Another advantage of the new measure is that, since only the average number of employees per establishment is taken into account for our classification purpose, it helps reduce the problem associated with the use of any official or unofficial definition of the size of manufacturing enterprises, which often vary from one country to another and from one organization to another.

In addition, the new measure of the relative share of the small enterprise sector versus the large enterprise sector (SEMSELE) differs from Ayyagari et al. measure of the small and medium enterprise sector (SMEOFF) in that SEMSELE only takes into account employment in the manufacturing sector. On the contrary, SMEOFF measures the employment share of the

[1] We fill in some of the missing annual values for an establishment category using linear interpolation method based on the data from that particular establishment category for the years where data are available.

SME sector in total official labor force – i.e. it measures SME share in the entire economy.

Another difference between the UNIDO-based and the Ayyagari et al. measures is that whereas the latter considers small and medium enterprises, the first measure takes into account only small and large enterprises and excludes those in between. Therefore, due to these differences, the two measures may differ from one another. In fact, the correlation between them is very low at 0.02.

Table 3.1: Sectoral classification based on median values of average number of employees per establishment in 1985

ISIC CODE	UNIDO Establishment Categories	Median Values of the Average Number of Employees (1985)	Sector Classification
332	Furniture, except metal	31	
390	Other manufactured products *	35	
323	Leather products	37	
331	Wood products, except furniture	37	**Small Enterprise Sectors**
381	Fabricated metal products	39	
342	Printing and publishing	42	
354	Misc. petroleum and coal products	43	
369	Other non-metallic mineral products	47	
385	Professional & scientific equipment	49	
356	Plastic products	49	
311	Food products	53	
322	Wearing apparel, except footwear	54	
382	Machinery, except electrical	56	
324	Footwear, except rubber or plastic	61	
361	Pottery, china, earthenware	68	
300	Total manufacturing	68	
352	Other chemicals	69	
355	Rubber products	74	
321	Textiles	83	
351	Industrial chemicals	88	
383	Machinery, electric	90	
362	Glass and products	90	
341	Paper and products	94	**Large Enterprise Sectors**
372	Non-ferrous metals	95	
384	Transport equipment	99	
313	Beverages	114	
371	Iron and steel	155	
314	Tobacco	244	
353	Petroleum refineries	333	

Notes: In the table we rank the 29 UNIDO establishment categories in accordance with their respective median values across countries of the average number of employees per establishment. We classify the 9 establishment categories with the largest median values as the large enterprise sector and the 9 categories with the lowest median values as the small enterprise sector. The remaining categories are excluded from our analysis since they may represent the medium enterprise sector, which is not considered in this study.

* We dropped the "Other manufactured products (ISIC code 390)" from the small enterprise sector because of its lack of specificity. In its place, we choose the "Plastic products (ISIC code 356)" to be included in the small enterprise sector.

Table 3.2: Variable Description and Sources

VARIABLES	DESCRIPTION	SOURCES
SMEOFF	Share of small and medium enterprises in manufacturing sector, when official country definition of SMEs is used.	Ayyagari, Beck and Demirguc-Kunt (2003) new database on SMEs
SEMSELE	The relative share of the small enterprise sector versus the large enterprise sector.	The data used to calculate SEMSELE are from UNIDO Industrial Statistics Database (2005)
LGDP	Log of real gross domestic per capita, measured in constant price, averaged over 1990-1999.	Summers, Heston and Aten's PWT version 6.1
$LGDP^2$	Square of the log of real gross domestic per capita.	Ibid.
LLY	Liquid liabilities – which equals to liquid liabilities of financial system (currency plus demand and interest-bearing liabilities of banks and nonblank financial intermediaries) divided by GDP.	Beck, Levine and Loayza (2000) financial dataset
BANK	The ratio of bank credits divided by bank credits plus central bank domestic assets.	Ibid.
PRIVATE	The ratio of private credits by deposit money banks and other financial institutions to GDP.	Ibid.
FINDEV	The measure of the level of financial development. It is the first principal component of BANK, LLY and PRIVATE.	Ibid.
KAOPEN	Index measure of capital account openness.	Chinn and Ito (2002)
SCH	Average years of schooling for the population age 15 and over.	Barro and Lee (2001) Educational Attainment Dataset, updated version
FDI	Ratio of the foreign direct investment to GDP.	World Bank World Development Indicators (2006)
ESEASIA	Dummy for East Asian countries.	Harvard University's Center for International Development (CID) geography dataset
EU	Dummy for European countries.	Ibid.
SAFRI	Dummy for Sub-Saharan African countries.	Ibid.
LATAM	Dummy for Latin American countries.	Ibid.
INST	A measure of institutional quality. It is the average of six different dimensions of governance: voice and accountability, political stability and violent, government effectiveness, regulatory quality, rule of law, and control of corruption.	Kaufmann et al. (1999)

VARIABLES	DESCRIPTION	SOURCES
STBUS	Costs of starting up a business as a percentage of GNI per capita.	World Bank Doing Business Data
LENTRCOSTPC	Log of the costs of obtaining a legal status for a firm to operate as a percentage of GDP per capita.	La Porta et al. (2002) Regulation of Entry dataset
MINCAP	Minimum capital required to start a business as a percentage of GNI per capita.	Ibid.
EMPLAW	Employment laws index measure of the protection of labor and employment laws.	La Porta et al. (2005) Regulation of Labor data
SSBENF	Social security laws index, measuring social security benefits.	Ibid.
CIVILR	Civil rights index which assesses the degree of protection of vulnerable groups against employment discrimination.	Ibid.
COLLR	Collective relations laws index measure of the protection of collective disputes and union power.	Ibid.
LEXPGDP	Log of exports as a percentage of GDP.	World Bank World Development Indicators (2006)
TARIFF	Average tariff on imported goods.	Ibid.
LND100CR	The 1994 proportion of a country's total area within 100km of the ocean or ocean navigable river.	Harvard University's CID geography dataset
POP100CR	1994 share of population within 100 km of ice-free coast/navigable river to total population.	Ibid.
AIRDIST	The closest distance in kilometers to a major port.	Ibid.
LANDLOCK	Dummy for landlocked countries.	Ibid.
OIL	Dummy for oil exporting (OPEC) countries.	OPEC website
INFRSTR	A measure of the quality of infrastructure. It is the first principal component of five measures of infrastructure: electric power transmission and distribution losses as a percentage of electricity production; percentage of roads that are paved; number of phone lines per 10,000 inhabitants; internet hosts per 1,000 inhabitants; and number of cars per 10,000 inhabitants.	World Bank World Development Indicators (2006).

VARIABLES	DESCRIPTION	SOURCES
SERVGDP	Share of the service sector in GDP.	World Bank World Development Indicators (2006).
URBPOP	Percentage of the urban population in total population.	Ibid.
POPGR	Population growth rate.	World Bank Macro Time Series Data.
INFL	Annual rate of inflation.	Ibid.
LRLP	Log of relative labor productivities in agriculture versus other sectors.	The data used to calculate RLP, namely the agricultural shares of output and labor are from the World Bank Development Indicators and FAOSTAT, respectively.

Table 3.3: Descriptive Statistics

Variable	Observation	Mean	Standard Deviation	Minimum	Maximum	25th Percentile	Median	75th Percentile
SEMSELE	41	53.76	10.85	36.67	82.51	45.75	50.73	60.89
SMEOFF	41	59.39	17.48	15.20	86.70	51.61	61.05	72.10
SMEOFF	76	51.44	22.7	4.59	86.70	33.60	58.54	69.32
LGDP	76	8.28	1.58	5.13	10.71	7.04	8.17	9.88
SCH	76	7.41	2.42	1.38	11.84	5.48	7.68	9.33
URBPOP	76	61.88	20.41	7.40	100	48.55	64.25	76.25
POPGR	76	1.08	1.06	-0.99	2.89	0.23	1.00	1.94
LRLP	76	1.09	0.59	-0.19	2.56	0.69	1.04	1.46
INFL	76	44.84	96.09	0.45	568.92	2.15	8.95	24.30
EU	76	0.23	0.42	0	1	0	0	0
SAFRI	76	0.13	0.34	0	1	0	0	0
LATAM	76	0.17	0.37	0	1	0	0	0
ESEASIA	76	0.13	0.34	0	1	0	0	0
OIL	76	0.02	0.16	0	1	0	0	0
INFRSTR	76	0.00	0.97	-1.31	2.00	-0.91	-0.29	0.98
SERVGDP	76	54.55	12.49	22.54	82.71	47.50	55.56	64.61
LEXPGDP	76	3.45	0.58	2.16	5.21	3.17	3.51	3.73
TARIFF	76	11.65	6.26	0.32	29.73	6.91	10.60	15.58
LANDLOCK	76	0.18	0.39	0	1	0	0	0
LND100CR	73	54.94	35.77	0	100	21	63	90
POP100CR	73	63.9	33.43	0	100	38	71	95
AIRDIST	75	3268.26	2555.29	140	9280	1150	2702.64	5230
FINDEV	76	0.00	0.94	-1.40	2.90	-0.67	-0.31	0.65

KAOPEN	76	0.54	1.47	-1.58	3.36	-0.76	0.09	2.07
FDI	76	2.51	2.01	0.10	9.70	1.10	1.95	3.80
INST	76	0.41	0.96	-1.46	1.93	-0.41	0.30	1.29
STBUS	76	26.15	38.96	0	222.4	4.50	10.40	29.53
LETRCOST	76	-1.63	1.15	-4.60	1.20	-2.12	-1.65	-0.77
MINCAP	76	43.15	72.78	0	378.6	0	17.40	41.85
EMPLAW	76	0.49	0.18	0.14	0.82	0.36	0.48	0.65
SSBENF	76	0.59	0.20	0.08	0.87	0.46	0.67	0.75
CIVILR	76	0.68	0.12	0.23	0.93	0.59	0.72	0.78
COLLR	76	0.46	0.11	0.18	0.71	0.38	0.46	0.55

Notes: This table shows descriptive statistics of all variables used in the analysis.

Table 3.4: Correlations

	SEMSELE	SMEOFF	LGDP	SCH	URBPOP	POPGR	LRLP	INFL	EU	SAFRI	LATAM	ESEASIA	OIL	INFRSTR	SERVGDP	LEXPGDP
SEMSELE	1.00															
SMEOFF	0.02	1.00														
LGDP	-0.20	0.55	1.00													
SCH	-0.22	0.17	0.66	1.00												
URBPOP	-0.17	0.32	0.78	0.65	1.00											
POPGR	0.32	-0.05	-0.44	-0.64	-0.40	1.00										
LRLP	-0.07	-0.21	-0.51	-0.40	-0.65	0.22	1.00									
INFL	-0.24	-0.39	-0.24	0.06	-0.04	-0.20	0.20	1.00								
EU	-0.02	0.37	0.63	0.29	0.36	-0.32	-0.35	-0.24	1.00							
SAFRI	0.22	-0.32	-0.52	-0.55	-0.52	0.55	0.35	-0.10	-0.21	1.00						
LATAM	0.27	0.09	-0.16	-0.31	0.02	0.39	-0.04	-0.02	-0.24	-0.17	1.00					
ESEASIA	-0.25	0.33	0.10	-0.02	0.01	0.17	0.06	-0.08	-0.21	-0.15	-0.17	1.00				
OIL	0.19	-0.02	-0.21	-0.20	-0.19	0.18	-0.02	-0.03	-0.08	0.17	-0.07	0.17	1.00			
INFRSTR	-0.18	0.40	0.91	0.70	0.70	-0.49	-0.52	-0.22	0.70	-0.41	-0.36	-0.01	-0.19	1.00		
SERVGDP	-0.15	0.53	0.80	0.43	0.68	-0.19	-0.45	-0.37	0.48	-0.37	0.14	0.01	-0.30	0.70	1.00	
LEXPGDP	-0.12	-0.01	0.21	0.19	0.21	-0.12	-0.18	0.01	0.09	-0.14	-0.28	0.32	0.04	0.20	0.13	1.00
TARIFF	0.06	-0.25	-0.69	-0.50	-0.61	0.46	0.52	0.23	-0.48	0.47	0.01	0.15	0.29	-0.69	-0.60	-0.12
LANDLOCK	-0.15	-0.29	-0.19	0.04	-0.26	-0.07	0.37	0.30	-0.07	0.13	-0.20	-0.17	-0.07	-0.12	-0.27	0.07
LND100CR	-0.14	0.34	0.45	0.26	0.39	-0.34	-0.42	-0.14	0.26	-0.50	-0.00	0.28	-0.05	0.34	0.39	0.36
POP100CR	-0.09	0.43	0.57	0.33	0.44	-0.30	-0.51	-0.23	0.34	-0.53	-0.01	0.27	0.01	0.50	0.52	0.28
AIRDIST	0.35	-0.10	-0.46	-0.47	-0.32	0.63	0.33	0.06	-0.47	0.51	0.28	0.03	0.13	-0.54	-0.28	-0.22
FINDEV	-0.31	0.53	0.74	0.39	0.50	-0.09	-0.30	-0.37	0.44	-0.26	-0.23	0.46	-0.08	0.69	0.63	0.27
KAOPEN	-0.06	0.48	0.77	0.44	0.55	-0.11	-0.37	-0.40	0.54	-0.35	-0.09	0.18	-0.01	0.73	0.75	0.20
FDI	-0.06	0.05	0.04	0.07	0.20	0.02	-0.13	-0.02	-0.05	-0.22	0.04	0.24	0.01	-0.00	0.04	0.52
INST	-0.08	0.56	0.90	0.56	0.68	-0.33	-0.57	-0.43	0.64	-0.38	-0.17	0.05	-0.22	0.88	0.79	0.25
STBUS	0.18	-0.33	-0.62	-0.66	-0.60	0.56	0.28	-0.02	-0.28	0.56	0.14	-0.04	0.18	-0.54	-0.46	-0.28
LETRCOSTPC	0.01	-0.37	-0.66	-0.57	-0.55	0.29	0.38	0.15	-0.33	0.30	0.18	-0.02	0.25	-0.67	-0.51	-0.09
MINCAP	-0.25	-0.09	-0.11	0.11	-0.11	-0.01	0.12	0.12	-0.09	0.04	-0.17	0.11	0.02	-0.09	-0.22	-0.06
EMPLAW	-0.07	-0.14	-0.00	0.13	-0.01	-0.51	0.11	0.25	0.29	-0.30	-0.12	-0.25	-0.05	0.05	-0.08	-0.00
SSBENF	-0.12	0.33	0.68	0.68	0.64	-0.65	-0.53	-0.08	0.41	-0.63	-0.06	-0.12	-0.27	0.65	0.54	0.15
CIVILR	0.08	-0.16	-0.25	-0.06	-0.16	-0.25	0.06	0.15	-0.20	-0.08	0.14	-0.35	-0.18	-0.24	-0.15	-0.15
COLLR	-0.14	0.01	0.04	0.12	0.11	-0.37	0.08	0.17	0.11	-0.29	0.08	-0.07	-0.22	-0.03	0.03	-0.14

Table 3.4 (Continue):

	TARIFF	LANDLOCK	LND100R	POP100CR	AIRDIST	FINDEV	KAOPEN	FDI	INST	STBUS	LETRCOSTPC	MINCAP	EMPLAW	SSBENF	CIVILR	COLLR
TARIFF	1.00															
LANDLOCK	0.04	1.00														
LND100CR	-0.45	-0.29	1.00													
POP100CR	-0.52	-0.41	0.90	1.00												
AIRDIST	0.48	0.07	-0.51	-0.53	1.00											
FINDEV	-0.40	-0.15	0.34	0.43	-0.28	1.00										
KAOPEN	-0.56	-0.19	0.37	0.53	-0.31	0.77	1.00									
FDI	-0.03	0.04	0.13	0.07	0.07	0.17	0.13	1.00								
INST	-0.68	-0.22	0.45	0.59	-0.39	0.73	0.77	0.14	1.00							
STBUS	0.33	0.10	-0.29	-0.32	0.35	-0.36	-0.34	-0.21	-0.55	1.00						
LETRCOSTPC	0.47	0.13	-0.14	-0.29	0.14	-0.49	-0.51	0.00	-0.66	0.48	1.00					
MINCAP	0.13	0.07	0.06	0.00	-0.09	-0.15	-0.22	-0.21	-0.21	0.26	0.16	1.00				
EMPLAW	-0.13	0.05	0.06	0.03	-0.34	-0.25	-0.11	-0.11	-0.06	-0.10	0.15	0.10	1.00			
SSBENF	-0.59	-0.17	0.41	0.48	-0.56	0.34	0.36	-0.00	0.59	-0.62	-0.51	-0.06	0.25	1.00		
CIVILR	0.15	0.07	-0.01	-0.09	-0.17	-0.39	-0.38	-0.04	-0.24	-0.05	0.25	-0.01	0.21	0.13	1.00	
COLLR	-0.17	-0.02	0.10	0.06	-0.22	-0.13	-0.12	-0.11	-0.12	-0.14	0.21	0.06	0.48	0.22	0.13	1.00

Notes: This table presents the correlations among the variables used for analysis in this chapter.

Table 3.5: Country List (when SMEOFF is the dependent variable)

ALB Albania	KEN Kenya
ARG Argentina	KOR Korea
AUS Australia	KGZ Kyrgyz Rep
AUT Austria	LVA Latvia
AZE Azerbaijan	LUX Luxembourg
BLR Belarus	MEX Mexico
BEL Belgium	NLD Netherlands
BRA Brazil	NZL New Zealand
BRN Brunei	NIC Nicaragua
BGR Bulgaria	NGA Nigeria
BDI Burundi	NOR Norway
CMR Cameroon	PAN Panama
CAN Canada	PER Peru
CHL Chile	PHL Philippines
COL Colombia	POL Poland
CRI Costa Rica	PRT Portugal
CIV Cote D'Ivoire	ROM Romania
HRV Croatia	RUS Russian
CZE Czech Republic	SGP Singapore
DNK Denmark	SVK Slovak Rep
ECU Ecuador	SVN Slovenia
SLV El Salvador	ZAF South Africa
EST Estonia	ESP Spain
FIN Finland	SWE Sweden
FRA France	CHE Switzerland
GEO Georgia	TWN Taiwan
DEU Germany	TJK Tajikistan
GHA Ghana	TZA Tanzania
GRC Greece	THA Thailand
GTM Guatemala	TUR Turkey
HND Honduras	UKR Ukraine
HKG Hong Kong	GBR United Kingdom
HUN Hungary	USA USA
ISL Iceland	VNM Vietnam
IDN Indonesia	YUG Yugoslavia
IRL Ireland	ZMB Zambia
ITA Italy	ZWE Zimbabwe
JPN Japan	
KAZ Kazakhstan	

Note: This table lists names and country codes for the 76 sample countries used for conducting empirical exercises, with the share of the small and medium enterprise sector (SMEOFF) as the dependent variable.

Table 3.6: BMA Results with SMEOFF as the Dependent Variable

Dependent Variable	SMEOFF	SMEOFF	SMEOFF	SMEOFF	SMEOFF	SMEOFF
Sample Countries	76	75	75	75	73	75
	(1)	(2)	(3)	(4)	(5) (Excl. Outliers)	(6) (MC3 Incl. Prob.)
LGDP	0.772 (+)	0.966 (+)	0.967 (+)	0.890 (+)	0.544 (+)	0.536
$LGDP^2$	0.463 (?)	0.865 (-)	0.871 (-)	0.877 (-)	0.511 (-)	0.437
SCH	0.058	0.063	0.053	0.085	0.305 (-)	0.112
URBPOP	0.360 (-)	0.130	0.126	0.140	0.022	0.102
POPGR	0.043	0.083	0.080	0.165	0.000	0.029
LRLP	0.047	0.009	0.000	0.117	0.062	0.046
INFL	0.737 (-)	0.905 (-)	0.909 (-)	0.054	0.098	0.077
EU	0.647 (+)	0.917 (+)	0.921 (+)	0.946 (+)	0.665 (+)	0.602
SAFRI	0.025	0.028	0.002	0.000	0.000	0.029
LATAM	0.833 (+)	0.359 (+)	0.321 (+)	0.402 (+)	0.598 (+)	0.402
ESEASIA	0.979 (+)	1.000 (+)	1.000 (+)	1.000 (+)	1.000 (+)	0.985
OIL	0.016	0.008	0.000	0.024	0.004	0.020
INFRSTR	0.103	0.041	0.039	0.060	0.045	0.429
SERVGDP	0.061	0.060	0.057	0.002	0.022	0.064
LEXPGDP		0.893 (-)	0.897 (-)	0.992 (-)	0.590 (-)	0.831
TARIFF		0.075	0.048	0.076	0.009	0.072
LANDLOCK		0.002	0.000	0.000	0.000	0.022
LND100CR		0.033	0.032	0.000	0.000	0.018
POP100CR		0.057	0.058	0.000	0.000	0.020
AIRDIST		0.520 (+)	0.542 (+)	0.197	0.271 (+)	0.038
FINDEV			0.054	0.000	0.000	0.040
KAOPEN			0.000	0.000	0.000	0.031
FDI			0.081	0.000	0.049	0.019
INST				0.974 (+)	0.950 (+)	0.928
STBUS				0.002	0.009	0.043
LETRCOST				0.000	0.000	0.025
MINCAP				0.000	0.000	0.015
EMPLAW				0.368 (-)	0.131	0.030
SSBENF				0.035	0.028	0.071
CIVILR				0.110	0.000	0.055
COLLR				0.000	0.010	0.022

Notes: The numbers reported in the table are the posterior probabilities of inclusion for each variable. The signs in the parentheses are the sign certainty indexes, which indicate the direction of the relationship between each variable and the dependent variable. It is based on the sum of posterior model probabilities for all the models in which a variable acts in a given direction (e.g. negative). For those numbers without the attached signs, it means that the direction of the relationship is uncertain. "MC3 Incl. Prob" is the inclusion probability for each variable when outlier-robust BMA is applied.

Table 3.7: Structures of the Top Ten Models and Their Posterior Probabilities as Suggested by BMA with SMEOFF as the Dependent Variable

Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
LGDP	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
$LGDP^2$	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SCH					✓					✓
POPGR	✓							✓		
EU	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
LATAM			✓				✓			
ESEASIA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
LRLP									✓	
LEXPGDP	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
AIRDIST						✓				
INST	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
EMPLAW		✓					✓		✓	✓
CIVILR								✓		
PMP	0.079	0.071	0.043	0.038	0.037	0.037	0.036	0.028	0.027	0.027

Notes: PMP stands for the posterior model probability, which is the probability that the model under consideration is the true model given the data. The sample consists of 76 developed and developing countries.

Table 3.8: OLS Estimates of the Top Ten Models from BMA with SMEOFF as the Dependent Variable

Regression Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Observation	76	76	76	76	76	75	76	76	76	76
CONSTANT	-179.231 (-2.891)	-139.808 (-2.465)	-78.375 (-1.397)	-105.622 (-1.876)	-118.455 (-2.137)	-149.776 (-2.630)	-111.624 (-1.928)	-192.857 (-3.106)	-158.074 (-2.766)	-143.449 (-2.550)
LGDP	66.573*** (4.317)	63.382*** (4.174)	41.868*** (2.800)	51.628*** (3.513)	54.949*** (3.798)	61.118*** (4.169)	53.369*** (3.354)	63.963*** (4.158)	66.417*** (4.393)	63.803*** (4.240)
$LGDP^2$	-4.218*** (-4.264)	-4.152*** (-4.173)	-2.778*** (-2.827)	-3.409*** (-3.525)	-3.423*** (-3.618)	-3.972*** (-4.121)	-3.508*** (-3.372)	-4.006*** (-4.044)	-4.347*** (-4.390)	-4.029*** (-4.075)
SCH					-2.071** (-2.025)					-1.577 (-1.513)
POPGR	4.954** (2.467)							5.991*** (2.843)		
EU	24.111*** (3.592)	29.548*** (3.910)	22.940*** (3.418)	21.504*** (3.131)	18.082** (2.610)	25.608*** (3.711)	29.383*** (3.954)	25.249*** (3.771)	30.012*** (4.017)	25.501*** (3.208)
LATAM			11.587** (2.220)				9.549* (1.823)			
ESEASIA	32.527*** (5.336)	34.317*** (5.671)	36.515*** (6.008)	35.179*** (5.658)	32.096*** (5.119)	36.552*** (6.011)	35.563*** (5.937)	34.977*** (5.588)	33.131*** (5.501)	32.124*** (5.209)
LRLP								5.835 (1.633)		
LEXPGDP	-12.261*** (-3.772)	-12.526*** (-3.836)	-10.981*** (-3.249)	-12.919*** (-3.849)	-11.982*** (-3.613)	-13.238*** (-3.879)	-10.995*** (-3.313)	-11.971*** (-3.710)	-12.200*** (-3.775)	-11.883*** (-3.642)
AIRDIST						0.002* (1.973)				
INST	14.766 *** (3.162)	14.199*** (2.995)	15.067*** (3.207)	16.052*** (3.339)	15.329*** (3.250)	15.878*** (3.355)	13.699*** (2.933)	14.213*** (3.062)	16.361*** (3.361)	13.982*** (2.976)
EMPLAW		-25.524** (-2.266)					-21.245* (-1.877)		-28.178** (-2.506)	-20.948* (-1.812)
CIVILR								25.729 (1.497)		
Adjusted-R^2	0.57	0.57	0.56	0.54	0 .56	0.57	0.58	0.58	0.58	0.57
Turning Point	2675	2065	1874	1944	3060	2194	2011	2932	1929	2746
Countries < TP	33	27	27	27	35	28	27	34	27	33

Notes: This table presents the Ordinary Least Square results for the top ten models selected through BMA exercises. Numbers in the parentheses are t-statistics. Turning Point is the dollar value of income where the share of the small and medium enterprise sector reverses course from rising to declining. Countries < TP is the number of countries that lie to the left of the turning point. *, **, *** indicate that the relationship is significant at 10, 5 and 1 percent level, respectively.

Table 3.9: Country List (when SEMSELE is the dependent variable)

ARG Argentina	NGA Nigeria
AUS Australia	NOR Norway
AUT Austria	PAK Pakistan
BGD Bangladesh	PAN Panama
BOL Bolivia	PER Peru
BGR Bulgaria	PHL Philippines
CMR Cameroon	POL Poland
CAN Canada	PRT Portugal
CHL Chile	ESP Spain
COL Colombia	LKA Sri Lanka
CRI Costa Rica	SWE Sweden
CYP Cyprus	THA Thailand
DNK Denmark	TUR Turkey
ECU Ecuador	GBR United Kingdom
EGY Egypt	VEN Venezuela
SLV El Salvador	ZWE Zimbabwe
FIN Finland	
FRA France	
GHA Ghana	
GRC Greece	
GTM Guatemala	
HUN Hungary	
ISL Iceland	
IND India	
IDN Indonesia	
IRN Iran	
IRL Ireland	
ISR Israel	
ITA Italy	
JPN Japan	
JOR Jordan	
KEN Kenya	
KOR Korea	
KWT Kuwait	
MAR Morocco	
NLD Netherlands	
NZL New Zealand	

Notes: This table lists names and country codes for the 53 sample countries used for conducting empirical exercises, with the UNIDO-based measure of the relative share of the small enterprise sector versus the large enterprise sector (SEMSELE) as the dependent variable.

Table 3.10: BMA Results with SEMSELE as the Dependent Variable

Dependent Variable	SEMSELE	SEMSELE	SEMSELE	SEMSELE	SEMSELE
Sample Countries	53	53	53	53	53
	(1)	(2)	(3)	(4)	(5) (MC3 Incl. Prob.)
LGDP	0.024	0.017	0.017	0.012	0.129
$LGDP^2$	0.016	0.027	0.026	0.013	0.152
SCH	0.051	0.395 (-)	0.391 (-)	0.409 (-)	0.282
URBPOP	0.096	0.274 **(+)**	0.274 (+)	0.230 (+)	0.062
POPGR	0.157	0.052	0.035	0.020	0.039
LRLP	0.068	0.060	0.058	0.051	0.015
INFL	0.003	0.017	0.017	0.006	0.026
EU	0.035	0.018	0.018	0.036	0.119
SAFRI	0.250 (+)	0.046	0.046	0.077	0.027
LATAM	0.313 (+)	0.249 **(+)**	0.240 (+)	0.337 (+)	0.298
ESEASIA	0.067	0.040	0.040	0.027	0.028
OIL	0.246 (+)	0.039	0.025	0.012	
INFRSTR	0.017	0.045	0.044	0.049	0.146
SERVGDP	0.201 (+)	0.295 (+)	0.294 (+)	0.240 (+)	0.342
LEXPGDP		0.937 (+)	0.933 (+)	0.940 (+)	0.814
LANDLOCK		0.053	0.053	0.039	0.131
LND100CR		0.013	0.013	0.009	
POP100CR		0.081	0.080	0.020	
AIRDIST		0.759 (+)	0.756 (+)	0.712 (+)	0.499
FINDEV			0.007	0.000	0.035
KAOPEN			0.025	0.000	0.035
FDI			0.017	0.010	0.031
INST			0.081	0.090	0.041
STBUS				0.007	0.187
LETRCOST				0.000	0.033
MINCAP				0.000	0.018
EMPLAW				0.008	0.017
SSBENF				0.001	
CIVILR				0.242 (-)	0.271
COLLR				0.018	

Notes: The numbers reported in the table are the posterior probabilities of inclusion for each variable. The signs in the parentheses are the sign certainty indexes, which indicate the direction of the relationship between each variable and the dependent variable. It is based on the sum of posterior model probabilities for all the models in which a variable acts in a given direction (e.g. negative). For those numbers without the attached signs, it means that the direction of the relationship is uncertain. "MC3 Incl. Prob" is the inclusion probability for each variable when outlier-robust BMA is applied. Due to program's limitation on the number of explanatory variables that can be included at any one time, some of the explanatory variables used in column 4 are excluded from MC3 exercise in column 5. These excluded variables are selected because they consistently have very little explanatory powers as shown in columns 1-4.

Table 3.11: Structures of the Top Ten Models and Their Posterior Probabilities as Suggested by BMA with SEMSELE as the Dependent Variable

Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
SCH				√		√		√		
URBPOP				√		√				
LATAM		√	√		√		√			
SERVGDP				√				√	√	
LEXPGDP	√	√	√	√	√	√	√	√	√	√
AIRDIST	√	√		√		√	√	√	√	√
CIVILR					√		√			√
PMP	0.081	0.045	0.045	0.040	0.035	0.030	0.029	0.027	0.025	0.024

Notes: PMP stands for the posterior model probability.
The sample consists of 53 developed and developing countries.

Table 3.12: OLS Estimates of the Top Ten Models from BMA with SMEOFF as the Dependent Variable

Regression Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Observation	53	53	53	53	53	53	53	53	53	53
CONSTANT	16.362	13.697	22.438	-7.418	38.474	11.078	29.032	-3.005	4.620	27.694
	(1.383)	(1.165)	(2.008)	(-0.500)	(2.744)	(0.912)	(1.985)	(-0.198)	(0.306)	(1.838)
SCH				-2.629***		-1.739**		-1.622*		
				(-2.827)		(-2.048)		(-1.967)		
URBPOP				0.189**		0.209**				
				(2.097)		(2.264)				
LATAM		6.237	8.265**		9.887**		7.870**			
		(1.627)	(2.180)		(2.594)		(2.029)			
SERVGDP				0.387**				0.431**	0.195	
				(2.047)				(2.215)	(1.237)	
LEXPGDP	10.062***	10.791***	9.352**	12.525***	8.919***	11.479***	10.289***	11.911***	10.016***	9.572***
	(3.003)	(3.244)	(2.805)	(3.820)	(2.729)	(3.432)	(3.140)	(3.525)	(3.005)	(2.849)
AIRDIST	0.002**	0.001*		0.002***		0.002**	0.001*	0.002**	0.002***	0.002**
	(2.448)	(1.953)		(2.765)		(2.376)	(1.841)	(2.529)	(2.732)	(2.440)
CIVILR					-22.747*		-20.856*			-14.739
					(-1.824)		(-1.707)			(-1.207)
Adjusted-R^2	0.16	0.19	0.14	0.27	0.18	0.22	0.22	0.21	0.17	0 .17

Notes: This table presents the Ordinary Least Square results for the top ten models selected through BMA exercises. Numbers in the parentheses are t-statistics. *, **, *** indicate that the relationship is significant at 10, 5 and 1 percent level, respectively.

Table 3.13: List of missing data for average number of employees per establishment

Country	Numbers of Missing Data	Years of Missing Data
Algeria	16	1981-1996
Argentina	2	1991, 1992
Australia	4	1993-1996
Austria	2	1995, 1996
Bangladesh	3	1993, 1994, 1996
Belgium	12	1985-1996
Bolivia	3	1970-1972
Brazil	11	1986-1996
Bulgaria	0	..
Cameroon	5	1973, 1985-1988
Canada	0	..
Chile	0	..
Colombia	0	..
Costa Rica	14	1970-1983
Cyprus	0	..
Denmark	0	..
Ecuador	0	..
Egypt	0	..
El Salvador	6	1986-1991
Finland	0	..
France	1	1996
Ghana	6	1988-1992, 1996
Greece	3	1994-1996
Guatemala	4	1989, 1990
Hungary	0	..
Iceland	0	..
India	0	..
Indonesia	4	1970-1973
Iran	1	1978
Ireland	0	..
Israel	2	1995-1996
Italy	0	..
Japan	0	..
Jordan	3	1970, 1972, 1973
Kenya	1	1996
Korea, Republic of	0	..

Kuwait	0	..
Malaysia	0	..
Malta	0	..
Mauritius	0	..
Morocco	6	1970-1975
Netherlands	0	..
New Zealand	0	..
Nicaragua	11	1986-1996
Nigeria	6	1979, 1986-1990
Norway	4	1993-1996
Pakistan	4	1992-1995
Panama	1	1995
Peru	10	1970-1978, 1993
Philippines	0	..
Poland	3	1994-1996
Portugal	1	1970
Singapore	0	..
Spain	0	..
Sri Lanka	0	..
Sweden	0	..
Thailand	11	1972-1973, 1978, 1980, 1981, 1983, 1985, 1987, 1992, 1995-1996
Turkey	0	..
United Kingdom	0	..
Venezuela	0	..
Zimbabwe	0	..

Notes: This table displays countries and years where the data for average number of employees per establishment are missing. “Numbers of Missing Data” indicates the number of years when data is (are) not available, while “Years of Missing Data” shows in what year(s) the missing data is (are).

Table 3.14: Descriptive Statistics of Average Number of Employees per Establishments in 1985

Countries	Observation	Mean	S.D.	Min	Max	25th Quartile	Median	75th Quartile
Argentina	29	45	124	3	670	10	14	28
Australia	29	89	157	14	833	29	47	67
Austria	29	115	82	8	405	63	89	156
Bangladesh	28	140	137	12	595	40	89	197
Bolivia	27	79	87	10	412	34	60	94
Brazil	23	40	28	1	119	18	34	63
Bulagaria	27	1127	1365	297	7000	507	656	908
Canada	29	90	82	16	305	32	59	102
Chile	29	193	130	66	583	129	144	186
Colombia	29	117	175	33	991	49	69	115
Costarica	28	68	156	5	848	15	27	64
Croatia	29	244	141	74	688	145	204	303
Cyprus	25	21	40	1	183	5	8	13
Denmark	29	72	46	19	205	39	60	89
Ecuador	29	84	67	31	335	50	68	83
Egypt	29	459	584	55	2950	103	263	470
Elsalvador	26	71	50	10	197	42	62	79
Fiji	17	18	11	4	50	10	17	21
Finland	29	145	252	33	1400	51	69	132
Greece	29	62	55	15	262	32	43	60
Guatemala	29	47	44	11	207	18	31	56
Honduras	27	55	43	8	186	26	39	71
Hungary	28	1109	1050	269	5500	500	844	1219
Iceland	27	32	79	1	360	3	6	14
India	29	102	188	20	1060	47	58	98
Indonesia	26	150	102	43	520	76	119	198
Iran	29	415	1380	32	7518	53	113	184
Ireland	27	51	43	13	238	25	45	59
Israel	27	58	54	11	233	25	41	62
Italy	25	128	126	23	571	63	91	128
Jamaica	22	54	66	14	326	25	36	44
Japan	29	45	67	10	362	17	21	48
Jordan	21	156	610	0	2812	3	8	28
Kenya	26	360	399	123	1988	195	235	317
Korea	29	83	99	22	518	38	54	77
Kuwait	23	282	1052	3	5093	12	47	81
Malawi	21	289	265	31	887	89	176	412
Malaysia	29	111	91	27	414	62	82	140
Malta	25	34	35	2	130	10	18	47
Mauritius	26	76	70	10	355	38	56	88
Mexico	27	398	275	113	1210	198	303	549
Morroco	18	73	56	22	272	38	63	77

Mozambique	27	196	211	20	870	80	110	200
Nepal	21	41	35	4	129	15	32	47
Netherlands	27	162	204	29	1000	52	83	161
Newzealand	29	21	17	3	74	11	16	24
Nicaragua	28	96	71	7	295	40	81	140
Nigeria	27	216	233	0	1000	70	143	203
Norway	29	85	83	21	299	31	44	86
Pakistan	29	149	185	34	967	63	92	162
Panama	26	42	25	10	117	24	39	49
Peru	29	41	60	10	259	15	21	36
Phillipine	29	154	135	40	575	71	110	189
Poland	29	990	856	333	4103	447	691	962
Portugal	29	94	135	0	740	36	54	100
Singapore	27	78	70	18	291	38	49	99
Spain	29	68	165	3	858	10	18	38
Srilanka	27	207	363	21	1811	47	84	209
Sweden	29	123	108	31	469	46	89	148
Tanzania	23	199	334	0	1673	59	119	239
Thailand	28	145	171	18	689	54	70	160
Turkey	29	219	209	60	925	101	147	262
UK	29	80	136	9	750	27	49	68
Venezuela	29	79	96	14	533	30	45	91
Zimbabwe	27	143	111	8	544	74	113	167

Table 3.15: Descriptive Statistics of Average Number of Employees per Establishments in 1981

Countries	Observation	Mean	S.D.	Min	Max	25th Quartile	Median	75th Quartile
Argentina	0	.	.	.	.	.	.	.
Australia	29	99	163	16	833	31	52	75
Austria	29	121	88	35	455	67	91	149
Bangladesh	28	127	116	11	509	37	84	215
Bolivia	28	56	94	3	366	16	26	34
Brazil	0	.	.	.	.	.	.	.
Bulagaria	26	1337	2131	319	11067	527	661	1008
Canada	29	105	104	17	367	38	59	110
Chile	29	198	117	85	553	120	162	222
Colombia	29	115	156	31	875	51	71	127
Costarica	0	.	.	.	.	.	.	.
Croatia	0	.	.	.	.	.	.	.
Cyprus	25	22	37	1	147	6	10	18
Denmark	29	74	49	14	200	38	56	88
Ecuador	29	87	80	23	412	48	65	87
Egypt	29	397	463	38	1750	78	214	477
Elsalvador	27	97	88	11	421	50	74	141
Fiji	18	22	18	0	72	11	19	25
Finland	29	155	238	36	1300	50	77	151
Greece	29	62	48	17	230	33	45	68
Guatemala	29	59	47	14	182	24	44	69
Honduras	0	.	.	.	.	.	.	.
Hungary	28	1518	1197	338	5500	612	1185	2147
Iceland	24	53	149	1	730	5	9	26
India	29	103	207	18	1165	42	63	89
Indonesia	26	149	98	37	432	78	125	182
Iran	29	492	1375	1	7131	33	111	232
Ireland	27	60	46	13	244	29	50	79
Israel	27	56	51	13	233	26	40	70
Italy	25	158	155	22	679	76	104	161
Jamaica	22	53	66	15	322	24	33	42
Japan	28	36	35	10	171	17	22	42
Jordan	21	146	522	0	2415	5	15	48
Kenya	27	370	238	129	1080	187	303	475
Korea	29	88	111	21	609	41	59	84
Kuwait	23	211	725	4	3518	12	36	71
Malawi	20	255	221	0	872	113	143	421
Malaysia	29	1409	3775	6	20431	116	408	810
Malta	25	33	35	2	139	11	20	43
Mauritius	25	79	80	0	398	36	55	83
Mexico	18	715	453	211	1936	379	585	992
Morroco	17	54	40	5	133	26	38	70

Mozambique	0	.	.	.	.	.	.	.
Nepal	0	.	.	.	.	.	.	.
Netherlands	27	115	204	2	1000	15	40	123
Newzealand	28	23	16	6	81	12	19	29
Nicaragua	27	89	85	28	447	47	70	104
Nigeria	25	353	287	51	918	98	272	607
Norway	29	84	81	20	300	31	43	94
Pakistan	29	146	190	26	933	60	90	160
Panama	27	52	44	0	237	30	46	60
Peru	29	47	68	9	333	15	22	46
Phillipine	29	70	107	4	425	10	30	73
Poland	29	1145	998	324	4973	522	834	1207
Portugal	29	97	135	0	733	36	55	90
Singapore	27	88	71	25	319	47	63	104
Spain	29	73	187	4	1000	12	20	40
Srilanka	29	241	512	28	2830	55	120	201
Sweden	29	126	122	32	562	47	84	148
Tanzania	23	192	232	0	1145	64	151	233
Thailand	27	129	118	17	490	57	99	131
Turkey	29	170	315	29	1440	46	79	127
UK	29	114	201	14	1100	35	74	103
Venezuela	29	76	75	17	394	28	55	101
Zimbabwe	27	144	112	10	562	79	125	170

Table 3.16: Descriptive Statistics of Average Number of Employees per Establishments in 1995

Countries	Observation	Mean	S.D.	Min	Max	25th Quartile	Median	75th Quartile
Argentina	29	22	34	3	184	7	12	19
Australia	21	43	80	6	364	11	20	35
Austria	24	57	73	10	343	14	29	77
Bangladesh	29	104	163	8	689	22	37	107
Bolivia	28	47	69	7	274	15	22	35
Brazil	19	148	60	83	368	120	139	166
Bulagaria	27	156	337	6	1472	11	26	81
Canada	28	79	62	19	277	36	57	112
Chile	29	202	109	93	653	145	170	236
Colombia	29	121	189	36	1015	56	71	96
Costarica	29	61	141	5	780	16	23	56
Croatia	29	72	156	4	828	19	28	52
Cyprus	26	15	29	2	143	4	6	10
Denmark	26	33	28	4	113	15	24	41
Ecuador	27	61	26	16	118	38	55	75
Egypt	29	332	413	45	2188	85	272	377
Elsalvador	28	93	79	8	256	31	64	149
Fiji	19	19	12	4	39	7	18	30
Finland	27	33	42	3	183	7	18	38
Greece	25	57	55	18	242	30	41	56
Guatemala	29	187	161	46	832	93	133	237
Honduras	26	159	130	16	496	66	112	209
Hungary	22	139	292	5	1395	26	37	119
Iceland	24	36	106	1	511	3	7	15
India	29	74	39	20	193	53	69	84
Indonesia	29	226	143	59	749	136	188	285
Iran	29	895	2426	1	12981	110	217	423
Ireland	0	.	.	.	.	.	.	.
Israel	21	42	25	15	113	22	38	49
Italy	28	33	89	2	477	5	9	25
Jamaica	20	70	43	21	179	40	48	109
Japan	29	42	53	11	292	17	26	50
Jordan	28	167	726	2	3863	6	14	42
Kenya	26	107	106	13	505	36	69	137
Korea	29	80	205	15	1107	22	28	48
Kuwait	25	283	1045	5	5277	15	41	81
Malawi	17	332	232	22	716	167	220	560
Malaysia	29	1548	4132	6	22453	104	501	1224
Malta	25	25	29	2	112	8	15	29
Mauritius	26	78	50	25	189	44	56	107
Mexico	27	563	313	107	1316	349	468	722

Morroco	27	107	97	21	479	49	74	119
Mozambique	27	144	152	15	660	54	90	136
Nepal	29	34	28	0	97	12	25	52
Netherlands	28	298	915	30	4944	58	88	169
Newzealand	29	34	66	2	285	8	14	24
Nicaragua	0	.	.	.	.	.	.	.
Nigeria	22	152	188	10	811	70	120	144
Norway	22	37	43	11	194	15	19	31
Pakistan	28	130	92	24	417	86	105	151
Panama	21	47	23	11	98	36	41	50
Peru	29	32	44	8	191	12	17	22
Phillipine	29	148	158	36	650	64	95	114
Poland	18	67	162	3	689	8	12	52
Portugal	29	149	638	6	3451	13	20	32
Singapore	27	90	66	25	270	39	64	127
Spain	28	36	59	6	244	10	15	30
Srilanka	28	89	121	10	615	31	48	95
Sweden	29	88	63	28	257	41	72	100
Tanzania	28	244	332	30	1712	66	146	300
Thailand	29	149	112	39	438	69	109	175
Turkey	29	154	203	43	1021	55	92	144
UK	28	56	101	8	563	21	33	58
Venezuela	29	82	91	17	500	33	49	97
Zimbabwe	27	229	190	22	880	106	181	264

Table 3.17: Sectoral classification based on means of the average number of employees per establishment in 1985

ISIC CODE	UNIDO Establishment Categories (1985)	Mean-of-Ratios	Sector Classification
390	Other manufactured products*	59	
332	Furniture, except metal	60	
323	Leather products	65	
342	Printing and publishing	70	
381	Fabricated metal products	72	**Small Enterprise Sectors**
356	Plastic products	76	
331	Wood products, except furniture	77	
385	Professional & scientific equipment	95	
322	Wearing apparel, except footwear	100	
382	Machinery, except electrical	104	
369	Other non-metallic mineral products	105	
311	Food products	106	
300	Total manufacturing	109	
352	Other chemicals	116	
361	Pottery, china, earthenware	150	
383	Machinery, electric	155	
341	Paper and products	164	
324	Footwear, except rubber or plastic	168	
313	Beverages	184	
384	Transport equipment	184	
321	Textiles	190	
354	Misc. petroleum and coal products	212	
351	Industrial chemicals	218	
372	Non-ferrous metals	222	**Large Enterprise Sectors**
355	Rubber products	223	
362	Glass and products	240	
371	Iron and steel	433	
314	Tobacco	472	
353	Petroleum refineries	718	

Notes: In the table we rank the 29 UNIDO establishment categories in accordance with their respective mean-of-ratios – i.e. mean values across countries of the average number of employees per establishment (measured in 1985). We classify the 9 establishment categories with the largest median values as the large enterprise sector and the 9 categories with the lowest median values as the small enterprise sector. The remaining categories are excluded from our analysis since they may represent the medium enterprise sector, which is not considered in this study.

* We dropped the "Other manufactured products (ISIC code 390)" from the small enterprise sector because of its lack of specificity. In its place, we choose the "Machinery, except electrical (ISIC code 382)" to be included in the small enterprise sector.

Table 3.18: Sectoral classification based on ratio-of-means of the average number of employees per establishment in 1985

ISIC CODE	UNIDO Establishment Categories (1985)	Ratio-of-Means	Sector Classification
361	Pottery, china, earthenware	11	
331	Wood products, except furniture	16	
332	Furniture, except metal	17	
342	Printing and publishing	23	
381	Fabricated metal products	23	**Small Enterprise Sectors**
390	Other manufactured products*	27	
323	Leather products	28	
356	Plastic products	31	
311	Food products	31	
322	Wearing apparel, except footwear	33	
385	Professional & scientific equipment	34	
369	Other non-metallic mineral products	35	
382	Machinery, except electrical	39	
300	Total manufacturing	40	
313	Beverages	41	
324	Footwear, except rubber or plastic	50	
354	Misc. petroleum and coal products	54	
341	Paper and products	58	
321	Textiles	63	
372	Non-ferrous metals	67	
383	Machinery, electric	76	
355	Rubber products	76	
352	Other chemicals	80	
314	Tobacco	81	**Large Enterprise Sectors**
351	Industrial chemicals	84	
362	Glass and products	93	
384	Transport equipment	107	
371	Iron and steel	117	
353	Petroleum refineries	384	

Notes: In the table we rank the 29 UNIDO establishment categories in accordance with their respective ratio-of-means – i.e. ratio of mean values across countries of the average number of employees per establishment (measured in 1985). We classify the 9 establishment categories with the largest median values as the large enterprise sector and the 9 categories with the lowest median values as the small enterprise sector. The remaining categories are excluded from our analysis since they may represent the medium enterprise sector, which is not considered in this study.

* We dropped the "Other manufactured products (ISIC code 390)" from the small enterprise sector because of its lack of specificity. In its place, we choose the "Wearing apparel, except footwear (ISIC code 322)" to be included in the small enterprise sector.

Table 3.19: Values of the median of average number of employees per establishment (median-of-ratios), mean-of-ratios (MOR) and ratio-of-means (ROM)

ISIC CODE	UNIDO Establishment Categories (in 1985)	Median-of-Ratios	Mean-of-Ratios	Ratio-of-Means
300	Total manufacturing	68	109	40
311	Food products	53	106	31
313	Beverages	114	184	41
314	Tobacco	244	472	81
321	Textiles	83	190	63
322	Wearing apparel, except footwear	54	100	33
323	Leather products	37	65	28
324	Footwear, except rubber or plastic	61	168	50
331	Wood products, except furniture	37	77	16
332	Furniture, except metal	31	60	17
341	Paper and products	94	164	58
342	Printing and publishing	42	70	23
351	Industrial chemicals	88	218	84
352	Other chemicals	69	116	80
353	Petroleum refineries	333	718	384
354	Misc. petroleum and coal products	43	212	54
355	Rubber products	74	223	76
356	Plastic products	49	76	31
361	Pottery, china, earthenware	68	150	11
362	Glass and products	90	240	93
369	Other non-metallic mineral products	47	105	35
371	Iron and steel	155	433	117
372	Non-ferrous metals	95	222	67
381	Fabricated metal products	39	72	23
382	Machinery, except electrical	56	104	39
383	Machinery, electric	90	155	76
384	Transport equipment	99	184	107
385	Professional & scientific equipment	49	95	34
390	Other manufactured products	35	59	27

Notes: This table displays together values of the median of average number of employees per establishment (median-of-ratios), mean values across countries of the average number of employees per establishment (mean-of-ratios, MOR) and ratio of mean values across countries of the average number of employees per establishment (ratio-of-means, ROM). All are measured in 1985.

Table 3.20: Comparing sectoral classifications based on median values of average number of employees per establishment (Median-of-Ratios), Mean-of-Ratios and Ratio-of-Means

ISIC CODE	UNIDO Establishment Categories	Median-of-Ratios	Mean-of-Ratios	Ratio-of-Means
332	Furniture, except metal	31	S	S
390	Other manufactured products *	35	S	S
323	Leather products	37	S	S
331	Wood products, except furniture	37	S	S
381	Fabricated metal products	39	S	S
342	Printing and publishing	42	S	S
354	Misc. petroleum and coal products	43	L	M
369	Other non-metallic mineral products	47	M	M
385	Professional & scientific equipment	49	S	M
356	Plastic products	49	S	S
311	Food products	53	M	S
322	Wearing apparel, except footwear	54	S	S
382	Machinery, except electrical	56	S	M
324	Footwear, except rubber or plastic	61	M	M
361	Pottery, china, earthenware	68	M	S
300	Total manufacturing	68	M	M
352	Other chemicals	69	M	L
355	Rubber products	74	L	L
321	Textiles	83	L	M
351	Industrial chemicals	88	L	L
383	Machinery, electric	90	M	L
362	Glass and products	90	L	L
341	Paper and products	94	M	M
372	Non-ferrous metals	95	L	M
384	Transport equipment	99	M	L
313	Beverages	114	M	M
371	Iron and steel	155	L	L
314	Tobacco	244	L	L
353	Petroleum refineries	333	L	L

Notes: This table compares sectoral classifications based on median values of average number of employees per establishment (Median-of-Ratios), mean values across countries of the average number of employees per establishment (mean-of-ratios, MOR) and ratio of mean values across countries of the average number of employees per establishment (ratio-of-means, ROM). All are measured in 1985. Arrangement of sectors in this table bases on median values of average number of employees. Letters "S", "M" and "L" represent small, medium and large enterprise sectors, respectively.

Figure 3.1: Relationship between income and the SME sector share

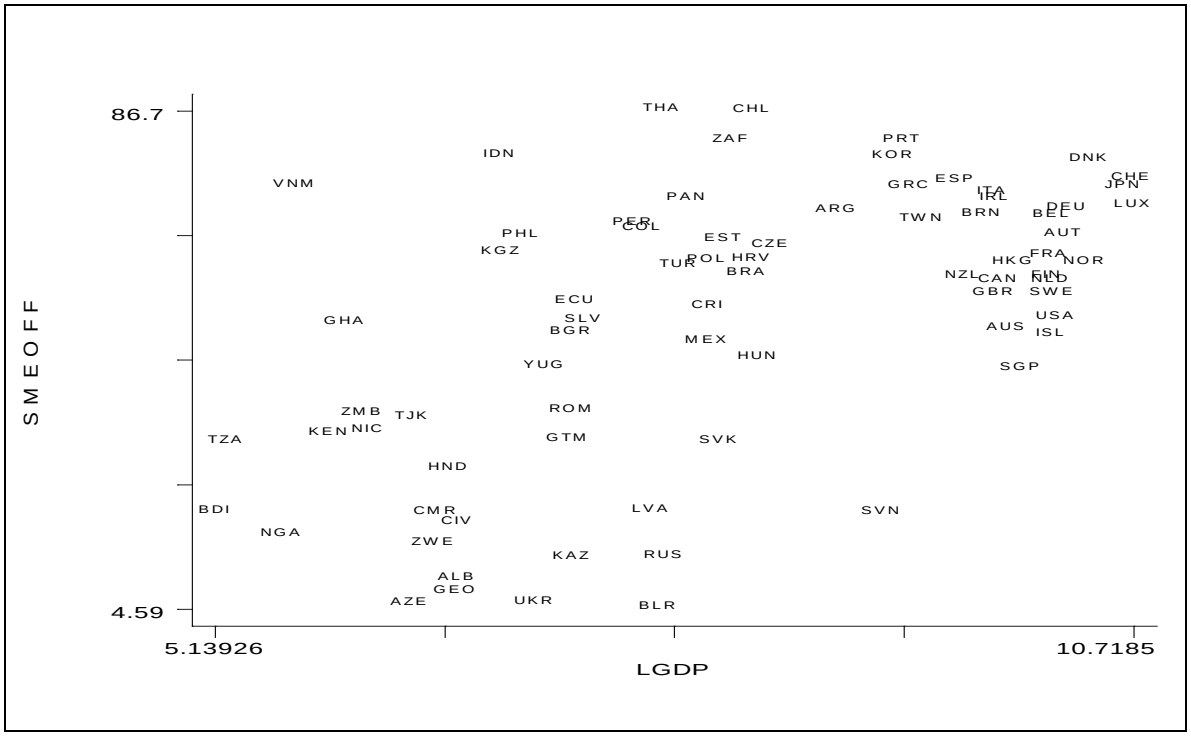


Figure 3.2: Relationship between human capital and the SME sector share

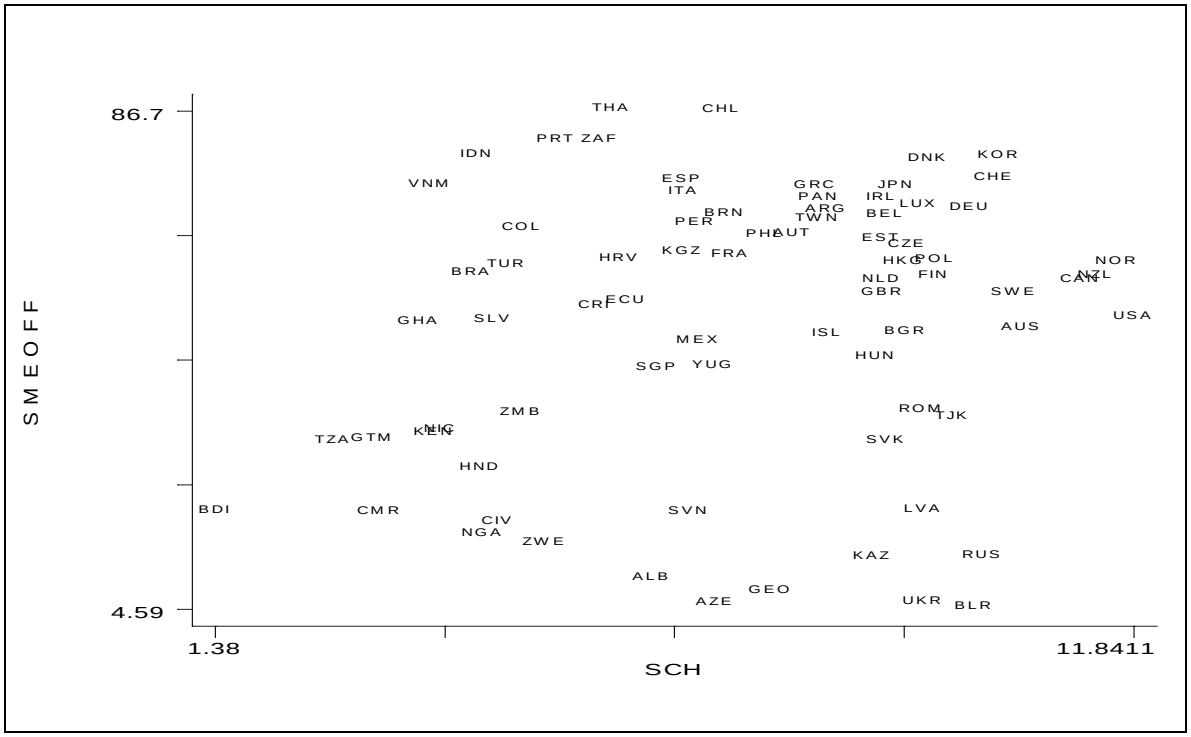


Figure 3.3: Relationship between the log of the ratio of exports to GDP and the SME sector share

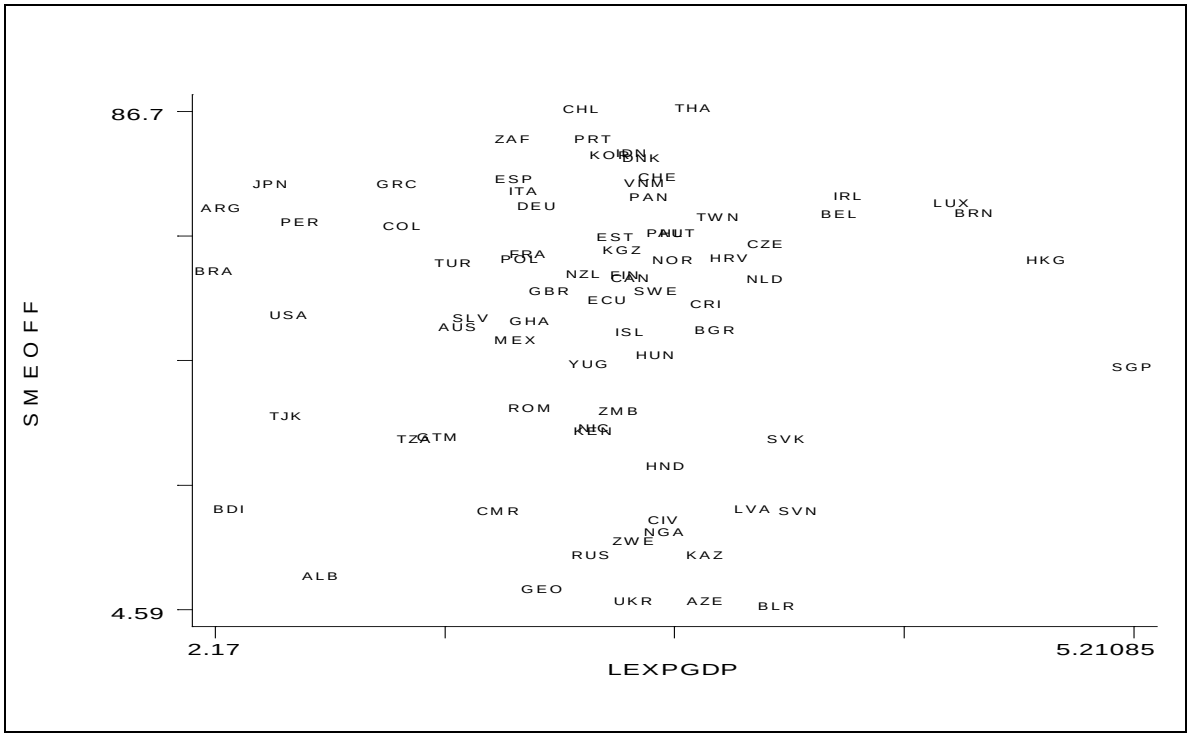


Figure 3.4: Relationship between institutional quality and the SME sector share

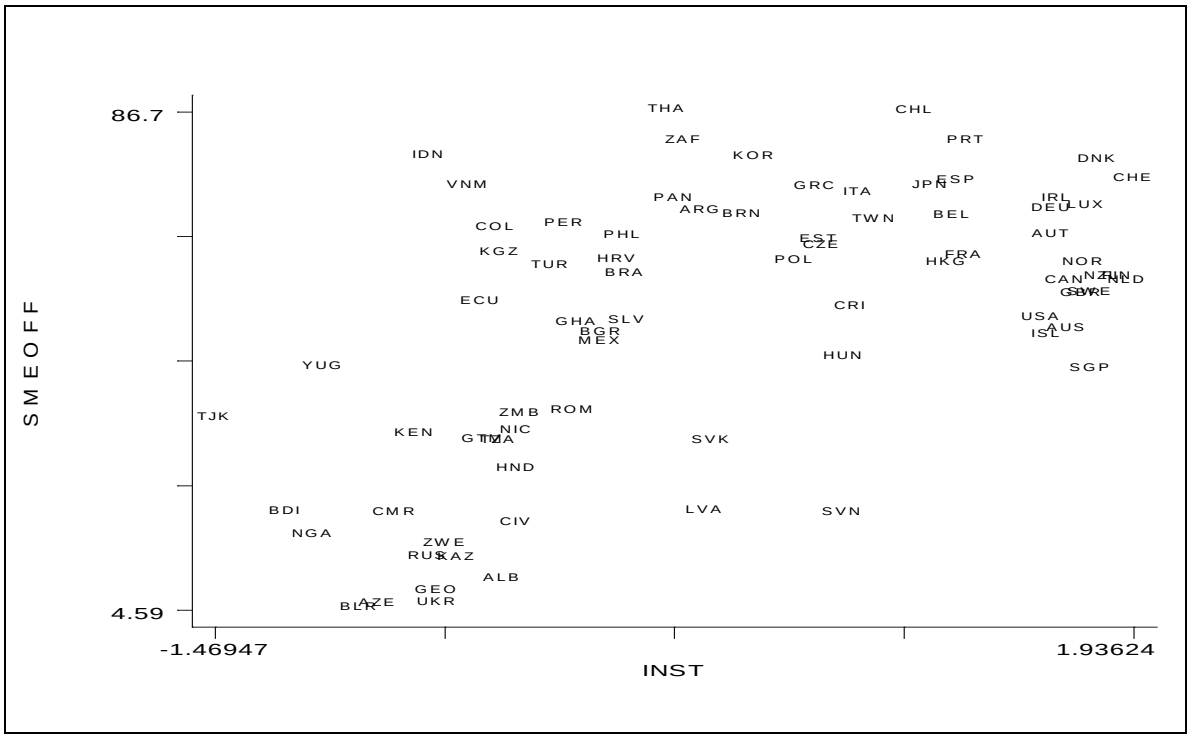


Figure 3.5: Relationship between human capital and the relative size of the small enterprise sector versus the large enterprise sector

Figure 3.6: Relationship between share of service sector in GDP and the relative size of the small enterprise sector versus the large enterprise sector

Figure 3.7: Relationship between urban population share of total population and the relative size of the small enterprise sector versus the large enterprise sector

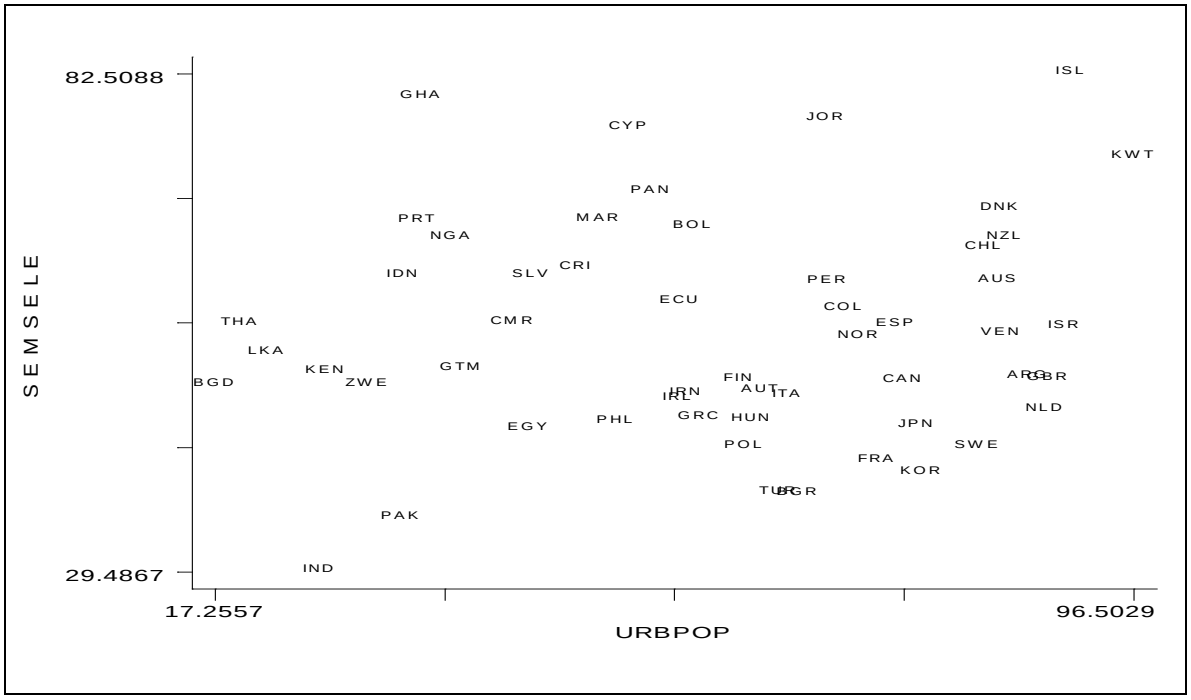


Figure 3.8: Relationship between log of the ratio of exports to GDP and the relative size of the small enterprise sector versus the large enterprise sector

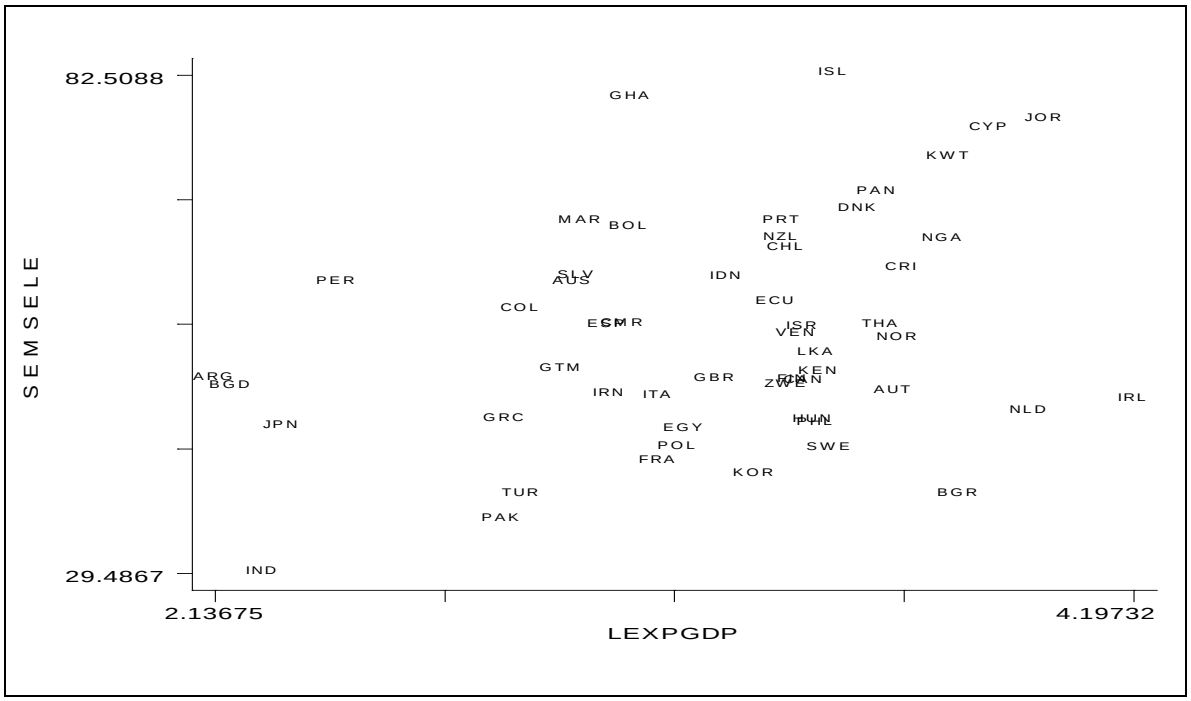


Figure 3.9: Relationship between air distance and the relative size of the small enterprise sector versus the large enterprise sector

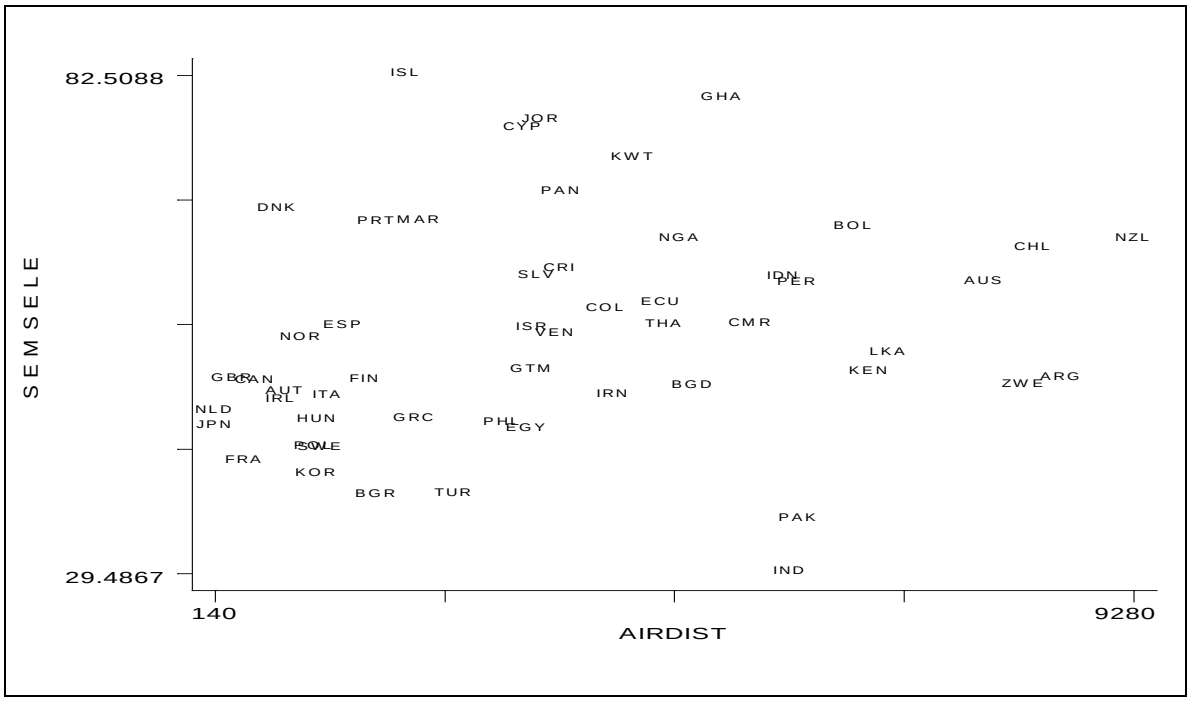


Figure 3.10: Density function plots for UNIDO 29 industrial sectors, in 1981

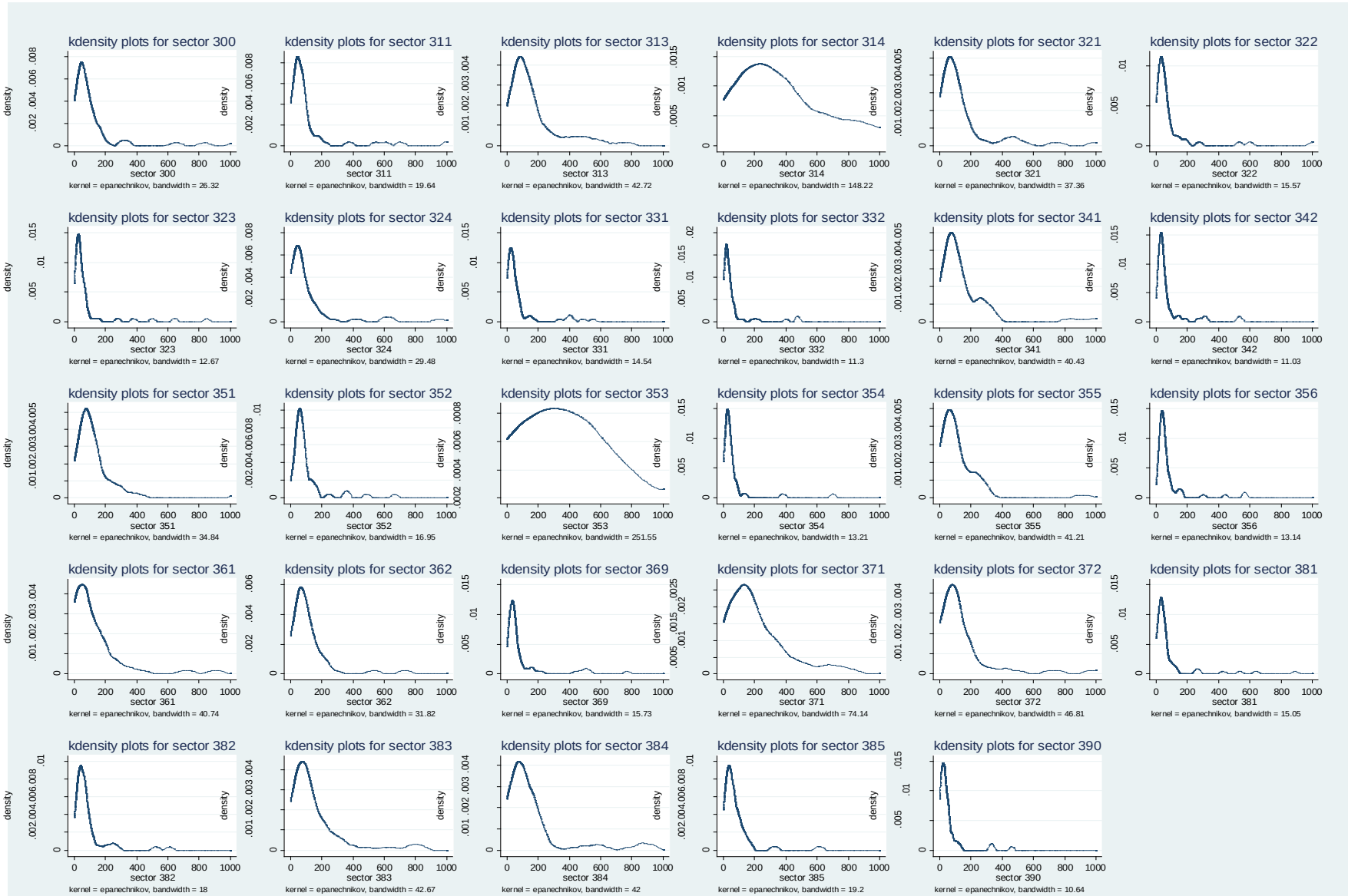


Figure 3.11: Density function plots for UNIDO 29 industrial sectors, in 1985

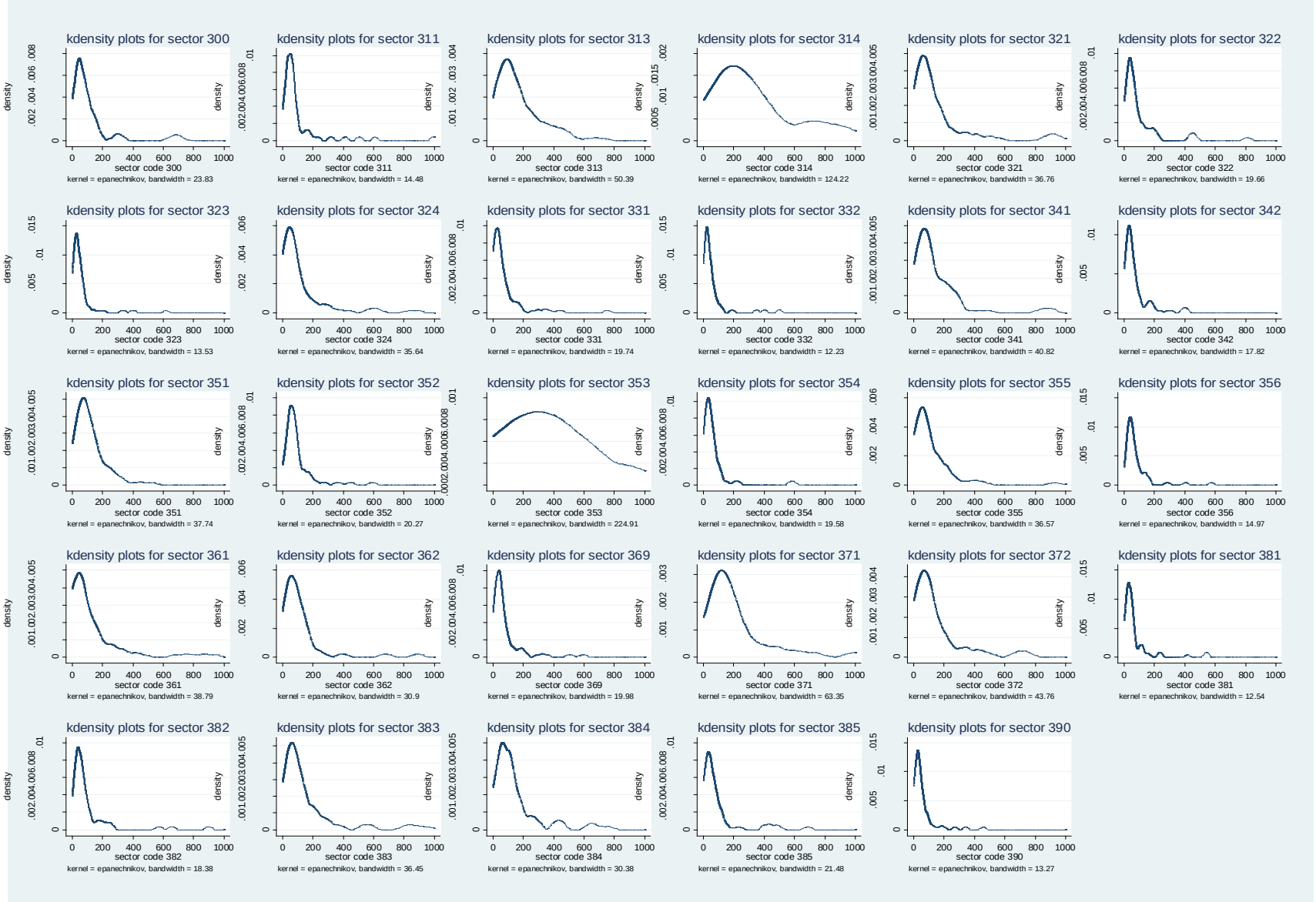


Figure 3.12: Density function plots for UNIDO 29 industrial sectors, in 1995

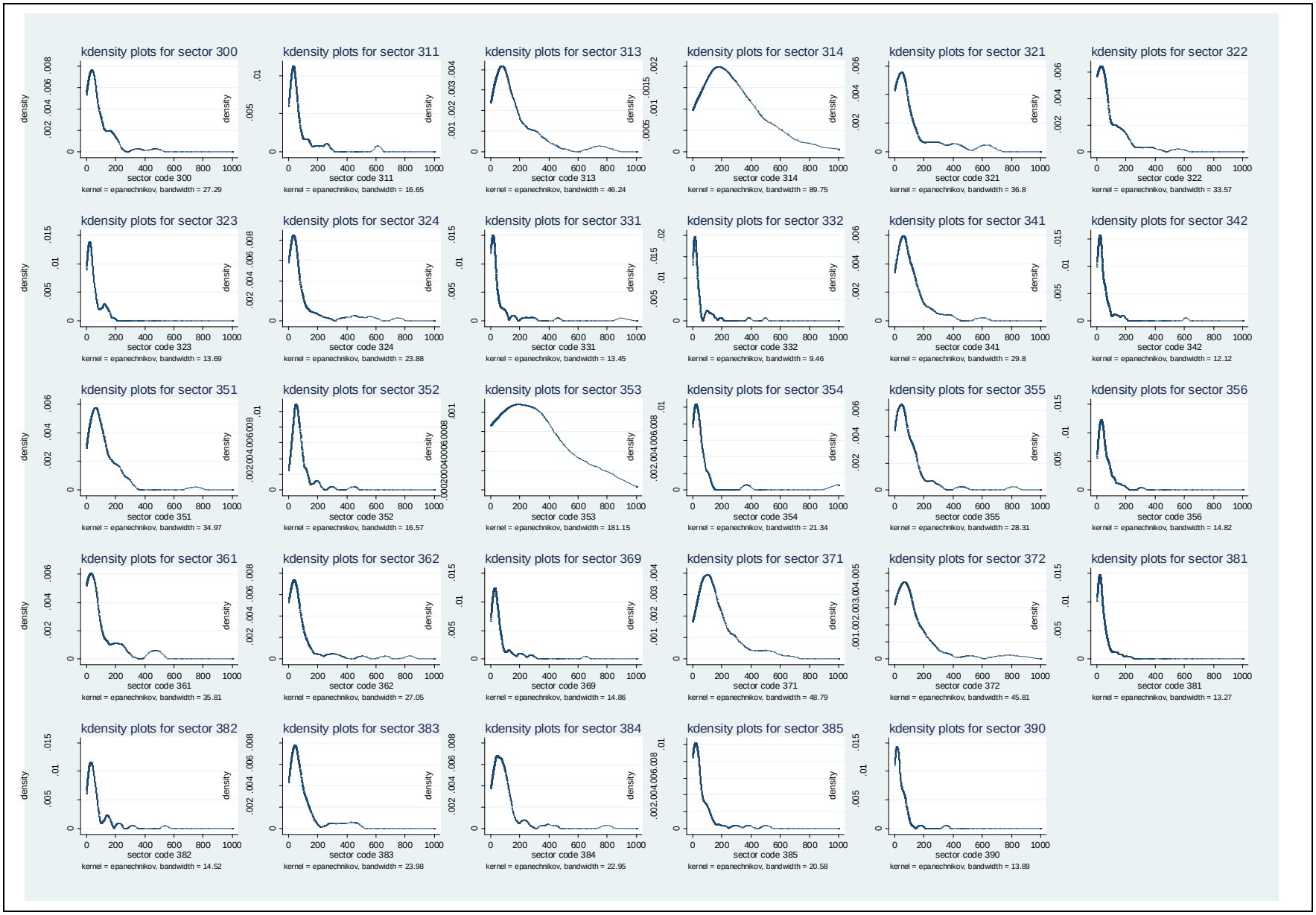


Figure 3.13a: Time series plots for the median values of the average number of employees per establishment

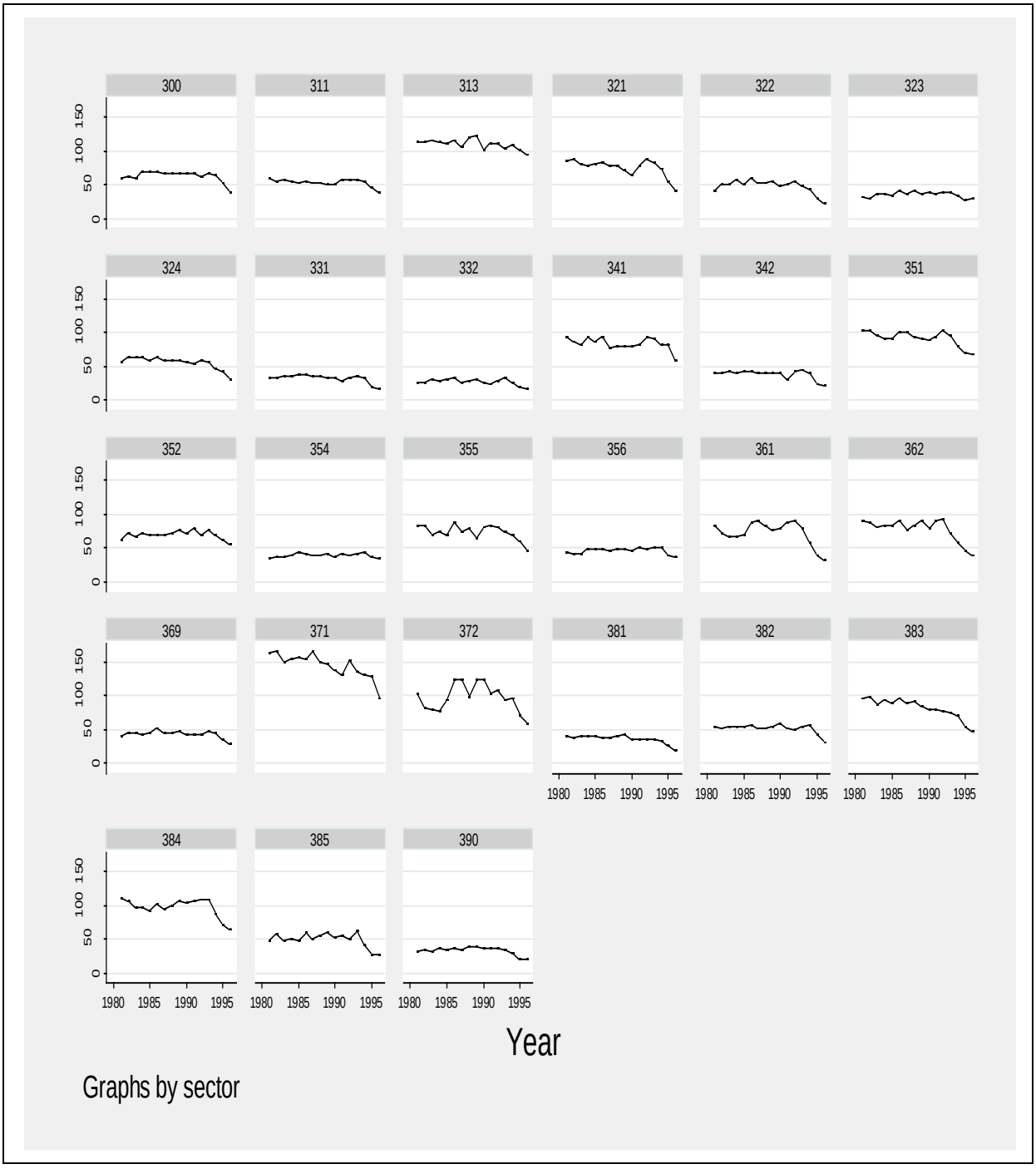
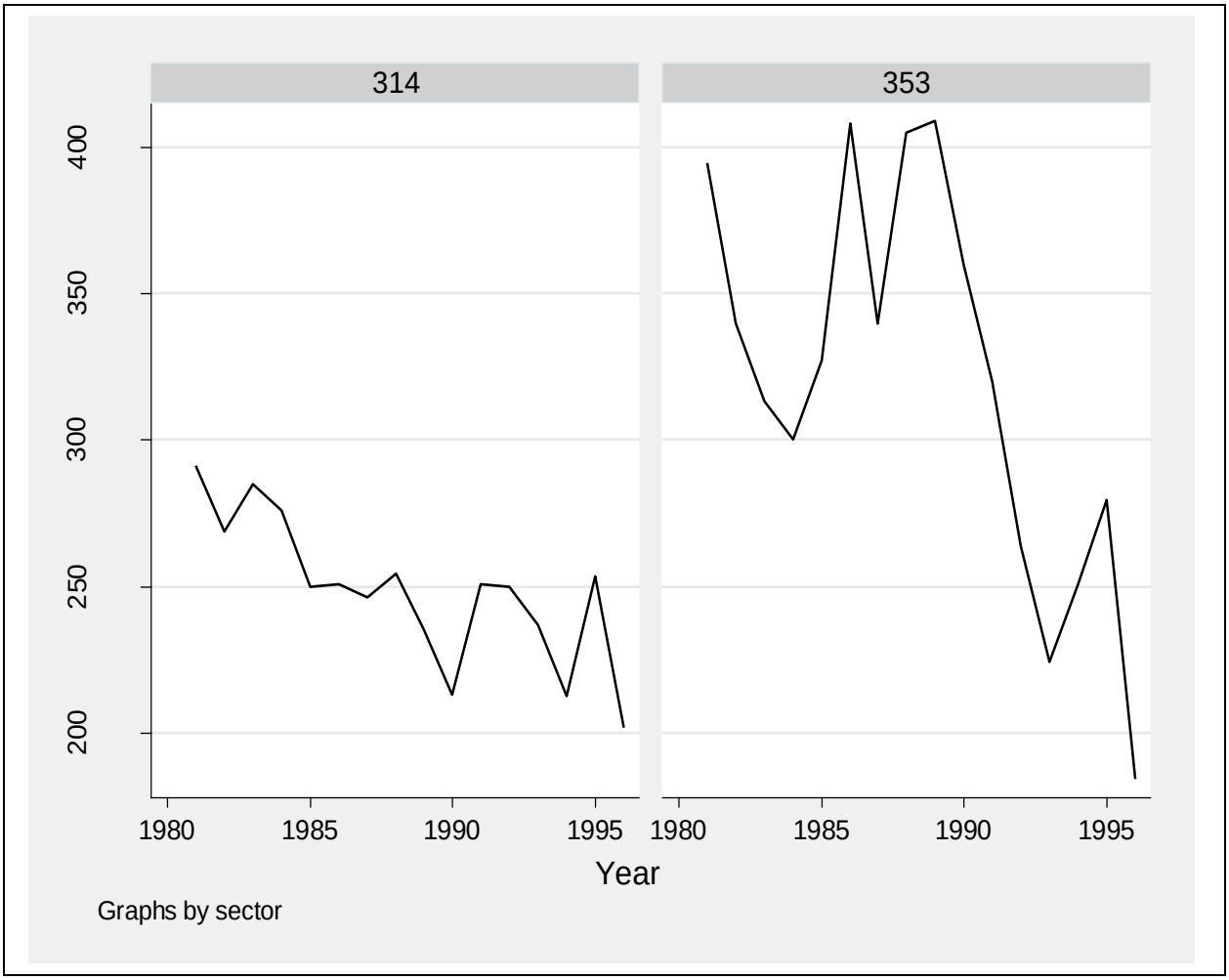


Figure 3.13b: Time series plots for the median values of the average number of employees per establishment



Notes: These figures show time series plots of the median values of the average number of employees per establishment for the 29 UNIDO industrial sectors between 1981 and 1996. Due to a large variation across sectors in the median values of the average number of employees per establishment, time series plots are broken into two separate graphs, with graph in Figure 3.13b shows time series plots for two industrial sectors with median values greater than 170. Figure 3.13a displays time series for the remaining 27 sectors. "MedianAVG" is the median values of the average number of employees per establishment. Numbers "300", "311", ect. are the industrial codes for the 29 UNIDO industrial sectors.

Chapter 4
The Effects of Financial Development and Liberalization on Small and Large Enterprise Sectors: A Panel Data Study

In developing countries, small enterprises may play a vital role in the economic growth and development process. Various studies by the World Bank suggest that small firms are more effective than large firms in promoting competition, innovation, and employment (World Bank 1994, 2002).

One of the main challenges commonly faced by small enterprises, however, is their inability to obtain access to credit.[37] Some research suggests this is a result of the underdeveloped state of the financial system.[38] An improvement in the level of financial development would ease such financial constraints because a well functioning financial market can serve as a direct source of capital and as a mechanism ensuring that investors have access to information about firms' activities.

Although there is evidence indicating the positive effect financial development has on the ability of small firms to grow, there is no consensus on the disproportionate effects financial development has on small relative to large firms. Some studies suggest that by lowering information and transaction costs in financial transactions, financial development would benefit small, less wealthy firms more than large firms.[39] On the contrary, other research argues that financial development disproportionately benefits large firms.[40]

---

[37] Some evidence that small firms are in a disadvantaged position in gaining access to inputs including credit is provided by Liedholm and Mead (1998) for several African and Caribbean countries and by Kinyanjui et al. (1997) for Nairobi's small and medium sized garment producers.

[38] Mayint (1971) argues that financial system underdevelopment engenders financial dualism in the sense that capital funds are usually provided to larger firms on easier terms than they are offered to small economic units. Moreover, the interest rate gap between funds given to large firms and those given to small firms tends to be larger in countries with less developed domestic credit markets.

[39] See, for example, Banerjee and Newman (1993); Galor and Zeira (1993); Aghion and Bolton (1997) and Jayaratne and Strahan (1998).

[40] Petersen and Rajan (1995) argues that creditors are more likely to finance credit-constrained firms when credits markets are concentrated since it would be easier for them to internalize the benefits of assisting firms. In this situation, it is also easier to form a mutually beneficial relationship between creditors and firms. However, financial development which fosters competition in the credit market would be detrimental to such

Similarly, there is no consensus among scholars about the disproportionate effects of financial liberalization on small and large enterprises. One the one hand, financial liberalization is said to benefit small, domestic firms by lowering borrowing costs as well as increasing the availability of credit to small firms.[41] On the other hand, international financial liberalization is seen to primarily benefit large, rich firms.[42]

In this chapter we seek to determine the potential disproportionate effects of financial development and financial liberalization on small firms and large firms. In particular, we examine whether financial development and financial liberalization allow the small enterprise sector – the sector dominated by small establishments – to grow faster than the large enterprise sector – the sector dominated by large establishments. If financial development disproportionately benefits the small enterprise sector, then we should see this sector expand faster than the large enterprise sector in countries with more developed financial systems. Likewise, if the small enterprise sector expands at a faster rate than the large enterprise sector in economies with more open financial systems, then we can suggest that financial liberalization enhances the growth of small enterprises more than large enterprises.[43]

The study covers 61 countries comprising both developed and developing countries over the period between 1970 to 1996. A major contribution of this study is that it employs a new measure of the relative size of the small enterprise sector in the economy derived from the United Nations Industrial Development Organization (UNIDO) dataset. This new measure provides a wider coverage of countries and over a longer time period compared to previous measures of the size of the small enterprise sector, and thus enables us to examine

---

a relationship and thus reduce the willingness of creditors to provide credit to small, credit-constrained firms. As a result, financial development may disproportionately hurt small firms.

[41] Alexander, Eun and Janakiramanan (1987), Domowitz, Glen and Madhavan (1998) and Hargis (2000) suggest that international financial liberalization boosts liquidity of domestic firms by transforming a segmented domestic equity market into an integrated market with high liquidity.

[42] Chowdhry and Nanda (1991), for example, argue that international liberalization of financial markets will induce a shift in the trading of international firms out of domestic market and into international markets. Such a shift reduces the liquidity available in the domestic market, which can have a negative effect on small firms that do not internationalize.

[43] One possibility that financial liberalization may allow large firms to gain access to credit from abroad, which, in turn, reduces competition for domestic credit. As a result, small firms may be able to gain better access to domestic credit and thus are able to grow and expand.

the impacts of financial development and financial liberalization on the small enterprise sector over time.

Methodologically, this study employs two econometric approaches. The first one is Generalized Method-of-Moments (GMM) dynamic panel data estimators. This method enables us to examine the variation in the relative size of the small enterprise sector using movements within countries over time, while allowing us to address the problems commonly associated with conventional cross-country growth regression studies such as measurement error and unobserved country-specific effects. The second method used in this analysis is a bias-corrected Least-Square Dummy Variable (LSDVC) estimator introduced by Kiviet (1995) and further developed in Kiviet (1999), Bun and Kiviet (2003) and Bruno (2005). The results from this method are compared with the results obtained from GMM methods.

The results from our estimation methods suggest a positive effect of the degree of capital account liberalization, as proxied by capital account openness, on the relative size of the small enterprise sector in the economy. Nevertheless, the strength of evidence depends greatly on the sample size (in terms of the number of time periods) used, with results become fragile when a restricted sample (fewer time periods) is used for estimations.

Moreover, the effect of domestic financial development is less clear. GMM estimates generally show a positive but fragile relationship between the indicator of financial development and the relative size of the small enterprise sector, while the LSDVC estimates indicate the opposite (i.e. a negative relationship).

The remainder of this paper is organized as follows. Section 2 will be a review of some of the previous research related to finance and the small enterprise sector, while in section 3 we describe the sample countries and data sources. Section 4 will present the methodologies employed in this study and results will be presented in section 5. Finally, section 6 will be the conclusion.

4.1 Previous Studies

It is often argued that one of the factors restricting the potential growth of small enterprises is their relative lack of access to formal credit compared to their larger counterparts. This phenomenon forces small firms either to resort to informal credit markets for credit sources or to give up any desired investment for growth.

Sleuwaegen and Goedhuy (2002) present the evidence that lack of access to inputs, especially credit, hurts small enterprises disproportionately. Using data for Ivory Coast, they show that restrained access to inputs results in a bi-modal firm size distribution with small firms growing slower and large firms growing faster than in the developed economies. Similarly, using firm-level survey data on over 4000 firms in 54 countries, Beck, Demirguc-Kunt, Maksimovic (2002) find that small firms are most adversely affected by financial constraints such as high interest rates, collateral requirements and lack of access to operations finance. They also argue that while such financial constraints may arise naturally in well-functioning markets, they are more severe in less developed financial systems.[44]

If small firms face more constraints in underdeveloped financial systems than large firms, then financial development, which eases such constraints and allow the markets to function more efficiently in providing credit, should disproportionately benefit small firms. Nabli and Nugent (1992) suggest that, in the case where transaction and information costs exist and financial constraints are binding, the more developed the credit market, the less the distribution by size should be biased against productive units of small size. In particular, they find that in less developed countries the removal of the sources of market failures in the financial market would likely shift the size distribution of manufacturing establishments in favour of the small and medium enterprise sector.

Guiso, Sapienza and Zingales (2002) investigate the impact of local financial development on firms' ability to grow. Using evidence across different regions in Italy, they find that local financial development enhances firms' growth by encouraging entry of new

[44] Besides financial constraints, Beck, Demirguc-Kunt, Maksimovic (2002) also examine the effects of legal and corruption constraints on firms' growth. Their results show that the extent to which these factors constrain a firm's growth depends very much on its size and that it is consistently the smallest firms that are most adversely affected by all these three constraints. In other words, firm growth is more affected by reported constraints in countries with undeveloped financial and legal systems and higher corruption.

firms, increasing competition and promoting growth. They also find that these effects are weaker for larger firms, which can more easily raise funds outside the local market.

Rajan and Zingales (1998) examine whether financial development has positive effects on industries that are dependent on external finance. They determine the industry's need for external finance – the difference between investments and internal cash generated from operations – using data on US firms.[45] By assuming that financial markets in the US are relatively frictionless, they can identify an industry's demand for external finance in a frictionless financial market. They also assume that the sector-specific demand for external finance is the same across countries. They then run a cross country regression using a sample of 43 countries (excluding the US) for the years 1980-1990 and find that industries that are more dependent on external finance grow faster in countries with more developed financial systems. Given this finding and the fact that small firms are generally more dependent on external finance than large firms (because of their limited abilities to generate sufficient internal capital), we can suggest that financial development could exert a disproportionately large positive effect on small firms.

In the same way, Beck, Demirguc-Kunt, Laeven and Levine (2004) find that industries that are composed of small firms grow faster in economies with well-developed financial systems.[46] Their results are robust to an array of sensitivity checks. In summary, the above studies provide evidence that financial development benefits industries dominated by small firms more than industries dominated by large firms.

Other research focuses on the impact of financial liberalization on small enterprises. Financial liberalization introduces greater efficiency and competition in the credit market, resulting in lower lending rates, which is very beneficial for small enterprises. Some theories imply that financial liberalization primarily benefits small firms because they are able to respond more quickly to changing economic conditions than their larger counterparts. Small firms can respond more flexibly under difficult and changing conditions because they do not depend heavily on infrastructure and because they can change their inputs and prod-

---

[45] They define a firm's dependence on external finance as the ratio of capital expenditures minus cash flow from operations divided by capital expenditures.

[46] Instead of considering only each industry's dependence on external finance, they also determine industries' composition of small firms relative to large firms. They define small firm share as the share of employment in firms with less than 20 employees in the US.

uct lines at relatively low cost due to the nature of their typically low levels of technology (Morawetz, 1974).

Uanto and Sanchez (1998) use a before and after study to determine the effects of financial liberalization on small-scale and micro enterprises in the food manufacturing sector in the Philippines. They find that financial liberalization had positive effects on the small enterprises' investment and financial positions through increases in their investment capital stock ratio as well as their profitability. Such improvements in their financial positions enhance their access to the formal credit market. In addition, financial liberalization allows formal financial institutions such as banks to expand their presence into many parts of the country and thus makes banking services more accessible even to non-traditional borrowers such as small and micro enterprises in the rural areas.

Using panel data on 394 firms in 13 developing countries for the years 1988-98, Laeven (2000) finds that financial liberalization affects small firms more than large firms by reducing imperfections in financial markets. He argues that small firms gain most from liberalization because the favoritism of preferential credit directed to large firms tends to disappear under liberalization. Small firms are financially constrained before liberalization begins but become less so after liberalization. He shows, however, that financial liberalization has little effect on the financing constraints of large firms because they have better access to preferential directed credit in the period before liberalization.

This study differs from previous research in four different aspects. First, this study examines the disproportionate effects of both the level of financial development and the degree of financial liberalization on the small and the large enterprise sectors. Second, while several previous studies focus on firm-level, country-specific evidence, which usually involves only a small number of countries, our study employs cross-country, cross-industry data covering a larger number of countries and over a longer period of time. Third, by using microeconomic data in their analysis, a number of previous studies focus on changes in the size of the small enterprise sector versus other sectors in the economy. In this study, only changes in the relative size of the small enterprise sector (the sector with low establishment sizes on average) with respect to the large enterprise sector are considered. Finally, because of the availability of annual data for a large number of countries, we can use panel data estimation methods which have several advantages over the time-series and cross-section

methods employed in previous studies. This will allow us to examine the changes in the relative size of the small enterprise sector within each country, controlling for individual country effects.

4.2 Sample Countries and Data

Our panel consists of annual data for 61 countries over the period between 1970 and 1996. The data used in this study are drawn from various sources including data on employment and number of manufacturing establishments from the United Nations Industrial Development Organization (UNIDO), the dataset on structure and development of the financial sector from Beck, Levine and Loayza (2000a, 2000b), and a measure of capital account liberalization obtained from Chinn and Ito (2002).

The dependent variable employed in this chapter is the UNIDO-based relative size of the small enterprise sector versus the large enterprise sector, SEMSELE. Ayyagari et al.'s measure of the share of the small and medium enterprise sector (SMEOFF) is not used because the available data for this variable do not contain the time-series variation required for conducting the empirical analysis.

We now turn our attention to financial factors. If financial development and financial liberalization disproportionately benefit small enterprises, we should see an expansion in the relative size of the sector dominated by small enterprises – i.e. those with low establishment size on average – as a result of improvements in the level of financial development and the degree of financial liberalization.

4.2.1 A Measure of Financial Development

A number of financial development measures have been constructed and employed in various studies recently.

This paper will use financial development indicators from the database on the structure and development of the financial sector compiled by Beck, Levine, and Loayza (2000a, 2000b). They use three quantitative indicators of the level of financial development to measure the functioning of the financial system. The first measure of financial development is

the commercial-central bank ratio (BANK), which equals the ratio of bank credit to bank credit plus central bank domestic assets. BANK measures the degree to which the central bank versus commercial banks is allocating credit. It can be used to indicate the relationship between the types of financial intermediaries that are conducting financial intermediation.

The second measure is called the liquid liabilities ratio (LLY) which measures the financial depth or the size of financial intermediaries in the economy. It equals liquid liabilities of the financial system (currency plus demand and interest-bearing liabilities of banks and non-bank financial intermediaries) divided by GDP. The final variable relates to private credit (PRIVATE), which indicates the amount of private credit extended by deposit money banks and other financial institutions as a ratio of GDP. It is calculated as private credit extended by deposit money banks and other financial institutions divided by gross domestic product.

In this paper, we construct an aggregate measure of the financial development using a principal components analysis. We denote this new aggregate variable as FINDEV. The advantage of using a combined measure rather than a single indicator is that it can help alleviate problems such as measurement error and outlier issues that might be associated with the use of a single indicator. In this case, our aggregate measure of financial development is the first principal component of the above three standardized measures.

Table 4.1 presents the proportion of all the principal components as well as the weights on each component. The first principal component accounts for 71% of the variation in these three indicators. By construction, FINDEV has a mean of zero. In terms of loadings on the individual variables, the aggregate indicator can be written as:

$$FINDEV_{it} = 0.513 * BANK_{it} + 0.592 * LLY_{it} + 0.622 * PRIVATE_{it}$$

where all the variables are standardized. The weights on these variables suggest that the variability of FINDEV is not driven predominantly by any single financial variable. In addition, they all have the expected positive signs.

4.2.2 A Measure of Financial Liberalization

There have been a number of different approaches employed in the empirical literature to quantify either the degree of financial repression or financial liberalization in an economy.[47] One of the commonly employed indicators of the degree of financial liberalization draws on the binary variables in the IMF categorical enumeration reported in *Annual Report on Exchange Arrangements and Exchange Restrictions (AREAER)*, which provides information on the extent and nature of restrictions on external accounts for a wide cross-section of countries.[48]

A possible weakness of the IMF's binary variables is that they do not indicate the degree of intensity of the controls being imposed on the financial system. They only show whether or not financial control is present.

Quinn (1997) created a composite index measure of financial regulation which assesses the many government policies regulating inward and outward financial transactions and allows comparison of the forms and intensity of that regulation across time and space. The index ranges from 0 to 14, with 14 representing the most open financial regime. The index is based on Quinn's coding of the qualitative information contained in IMF's AREAER pertaining to restrictions on current account (k_2) and capital account transactions (k_3), augmented by information regarding whether the country in question has entered into international agreements with international organizations such as the OECD and the European Union. His measures cover 64 countries for the period between 1950 and 1994. Quinn's index has a clear advantage over the IMF's variables in the sense that it measures the intensity of financial restrictions being applied in a country. However, as Chinn and Ito (2002) point out, one important shortcoming of the Quinn index is that while a complete tabulation for the OECD member exists, coverage for less developed countries is much less extensive, with values reported for only certain years (1958, 1973, 1982, and 1988).

---

[47] See Edison et al. (2002) for details on the comparison of different methods of quantifying capital controls.

[48] These binary variables are created based on a set of "on-off" classification, which include an indicator variable for the existence of multiple exchange rates (k_1); restrictions on current account (k_2); capital account transactions (k_3); and a variable indicating the requirement of the surrender of export proceeds (k_4). The most relevant capital controls are k_2 and k_3, which indicate restrictions on the current account and capital account, respectively.

A similar index of financial liberalization was introduced by Abiad and Mody (2003) for 35 countries over the 24-year period from 1973 to 1996. Drawing on available surveys of financial liberalization experiences, Abiad and Mody constructed this index as an annual aggregation of financial reform along six different policy dimensions including credit controls, interest rate controls, entry barriers and/or lack of pro-competition policies, regulations and security market restrictions, financial sector privatization, and restrictions on international financial transactions. Similar to the Quinn index, this new index of financial liberalization captures the changes in the intensity of financial repression over time. Nevertheless, the number of countries covered by the dataset is fairly small.

We follow Chinn and Ito (2002) in using an index measure of capital account openness which they constructed based on the AREAER binary series. In order to examine the effect of financial liberalization – rather than controls – they reverse the values of the AREAER binary variables such that the variable takes a value of zero when a restriction is in place and one when there is no restriction. The index is called KAOPEN and is the first standardized principal component of the IMF's binary variables k_1 (the existence of multiple exchange rates), k_2 (restrictions on current account), k_3 (restriction on capital account transactions) and k_4 (the requirement of the surrender of export proceeds).[49] This index measure has an advantage over the IMF variables, as it does not only show whether or not financial restrictions are in place, but it also indicates the intensity of such restrictions. In addition, this index covers a longer period of time compared to the Quinn index.

4.3 Framework and Methodology

The Generalized Method of Moments (GMM) estimator for dynamic panel data models will be used in the first part of our analysis.[50]

The GMM estimation method was designed to address issues such as unobserved country-specific effects and measurement errors which are commonly associated with growth regression studies. It has a number of advantages over the more traditionally used cross-

[49] See Chinn and Ito (2002) for a more detailed description of the calculation of the KAOPEN index.

[50] The GMM estimation method was proposed by Chamberlain (1984), Holtz-Eakin, Newey and Rosen (1988), Arellano and Bond (1991) and Arellano and Bover (1995).

section estimation method. First, a pure-cross section method rarely allows dynamic relationships to be investigated. In contrast, panel data allow us to examine how financial development over time within a country may have an effect on small enterprise growth and the relative size of this sector. Another advantage is that in panel data estimation we can control for unobserved country-specific effects and thus reduce biases in the estimated coefficients, whereas any unobserved country-specific effect in pure cross-section estimation would be incorporated into the error term, potentially causing biased coefficient estimates.

Our panel consists of data for 61 countries over the period 1970-1996. Table 4.2 provides a list of the countries used in this study. Data are averaged over non-overlapping, three-year periods, so there are nine observations per country (1970-1972; 1973-1975; ... ; 1994-1996). By averaging data over a period of time, the impact of measurement error and the likelihood that the results are driven by co-movements at very short horizons could be reduced. Our panel dataset is unbalanced, however.

We begin with a simple form of the dynamic panel data model, namely a first order autoregressive model, AR(1):

$$y_{it} = \alpha_1 y_{it-1} + \beta_1 x_{it-1} + \eta_i + \varepsilon_{it} \qquad -1 < \alpha_1 < 1;\ i = 1, ..., 61 \text{ and } t = 2, ..., 9 \quad (4.1)$$

where y_{it} is the measure of the relative size of the small enterprise sector, x_{it} is either a measure of financial development or of financial liberalization, η_i is an unobserved country-specific effect, ε_{it} is the error term and subscripts i and t represent country and time period, respectively.

To eliminate the country-specific effect, we can take first-differences of the above equation

$$\Delta y_{it} = \alpha_1 \Delta y_{it-1} + \beta_1 \Delta x_{it-1} + \Delta \varepsilon_{it} \qquad i = 1, ..., 61 \text{ and } t = 3, ..., 9 \qquad (4.2)$$

Under the assumption that the error term, ε, is not serially correlated and that the explanatory variables, x, are not correlated with future realizations of the error term (i.e.

the x's are weakly exogenous), a GMM dynamic panel estimator can use the following moment conditions:

$$E\left(y_{it-s}\,\Delta\varepsilon_{it}\right)=0 \qquad s\geq 2 \text{ and } t=3,...,9$$

$$E\left(x_{it-s}\,\Delta\varepsilon_{it}\right)=0 \qquad s\geq 2 \text{ and } t=3,...,9$$

We can estimate the first-differenced equation (4.2) using all available lagged levels of x_{it} and y_{it} dated $t-2$ and earlier. We also estimate the first-differenced equation with a restricted set of instruments in order to avoid potential overfitting biases which are sometimes associated with using all the available (linear) moment conditions. In this case, the reduced set of instruments uses no lags dated further back than $t-4$.

A fundamental drawback of the first-differenced GMM estimation method is that when the series are highly persistent over time, lagged levels of the series might be weak instruments for first differences. Arellano and Bover (1995) and Blundell and Bond (1998) suggest the use of a "system GMM" estimator, which combines equations in differences and equations in levels, to reduce the potential bias due to weak instruments.

In system GMM, the instruments for the equation in differences are the same as before. The instruments for the equation in levels (the untransformed equation) are the lagged first-differences of the corresponding variables dated $t-1$. These instruments are appropriate under an additional assumption that though there might be a correlation between the levels of the right-hand side variables and the country-specific effect in equation (4.1), there is no correlation between the differences of these variables and the country-specific effect.

The additional moment conditions are:

$$E\left[\Delta y_{it-1}\left(\eta_i+\varepsilon_{it}\right)\right]=0 \qquad t=3,...,9$$

$$E\left[\Delta x_{it-1}\left(\eta_i+\varepsilon_{it}\right)\right]=0 \qquad t=3,...,9$$

To check the validity of the instruments and thus the consistency of the GMM estimator, we examine two specification tests: the Sargan test of over-identifying restrictions and the test of no serial correlation in the error terms. The Sargan test of over-identifying restrictions tests the validity of the instruments by considering the sample analogue of the moment conditions used in the estimation process. Particularly, it compares the sample moment conditions with their population analogue.

The serial correlation tests are used to test whether the differenced error term is second-order serially correlated in both the first-differenced and the system GMM estimation methods. In other words, they are used to examine the null hypothesis of no first-order serial correlation and no second-order serial correlation respectively, in residuals in first-differences. When the errors in levels are serially uncorrelated, we can expect to find significant first-order serial correlation, but no significant second-order serial correlation in the first-differenced residuals. Note that, by construction, the differenced error term is probably first-order serially correlated even if the original error term is not.

In addition, to check for the possible bias of the GMM estimates, we compare our GMM results with the results obtained from OLS levels and Within Groups estimates. In a model like the AR (1) above, the OLS levels estimator is biased upwards in the presence of individual specific effects, and the Within Groups estimator is biased downwards.[51] Thus, a consistent estimate of α_1 should lie between the OLS levels and Within Groups estimates. If the result from the GMM estimate lies above the OLS levels estimate, then the GMM estimate is likely to be biased upwards. Similarly, if the GMM estimate is below the Within Groups estimate, then the GMM estimate is likely to be biased downwards.[52]

An important drawback associated with these estimators, however, is that their asymptotic properties depend on having a large number of cross-section units. One of the main problems in using these estimators in other contexts is that they may have poor finite sample properties in terms of bias and imprecision.

Kiviet (1995) uses asymptotic expansion techniques to approximate the small sample bias of the LSDV estimator and thus offers a method to correct the LSDV estimator for samples where N is small or only moderately large. The Monte Carlo evidence in Jud-

[51] For a more techinical discussion of this issue, see Bond (2002).

[52] See Nerlove (1999, 2000) for discussion of this issue in the context of empirical growth models.

son and Owen (1999) and Bun and Kiviet (2003) suggests that the bias-corrected Least Square Dummy Variable estimator (LSDVC) is more effective than both first-differenced GMM and system GMM in terms of bias and root mean square error (RMSE) for small or moderately large samples.[53] This finding was nevertheless limited to balanced panel data. Bruno (2004), by extending the bias approximation formulas in Bun and Kiviet (2003), was able to derive a bias approximation of various orders in dynamic unbalanced panels with a strictly exogenous selection rule. For these reasons, the results from the LSDVC estimation method will receive special attention in this study.

4.4 Empirical Results

In this section we present the findings from our dynamic panel data estimates. In summary, the results indicate a positive and significant impact of the degree of financial liberalization on the relative size of the small enterprise sector. On the other hand, the evidence that the level of domestic financial development matters is less clear. The results for our analysis of the effects of financial development and of financial liberalization in terms of capital account openness will be presented in Table 4.3 and Table 4.4, respectively.

For comparison purposes, we present results from OLS levels, the Within-Group, Differenced-GMM, System-GMM, and the bias-corrected Least Square Dummy Variable (LSDVC) estimates. As explained earlier, in the AR (1) model, the OLS estimate of the autoregressive parameter is biased upwards while the Within-Groups estimate is biased downward. It follows that a consistent estimate of the autoregressive parameter should typically lie somewhere in between the OLS levels and the Within-Group estimates. It is a simple indication of the presence of serious finite sample biases when particular estimates do not fall into this interval or are very close to the bounds. For the GMM estimation method, we examine the tests of no serial correlation as well as the Sargan test of the over-identifying restrictions.

A brief look at Figure 4.1 indicates a positive relationship between the relative size of the small enterprise sector and the level of financial development. Table 4.3 presents the

[53] A potential disadvantage of the LSDVC method is that it treats all explanatory variables other than the lagged dependent variable as strictly exogenous.

results of our empirical estimation examining the effects of financial development on the relative size of the small enterprise sector. These results indicate that the overall effect of financial development on the relative importance of small enterprises is ambiguous. The results from OLS levels, Within-Groups and LSDVC estimates in column (1), (2) and (8), respectively, suggest that financial development exerts a negative although weak effect on the relative size of the small enterprise sector. On the contrary, as shown in column (3) through (7), estimating the relationship using GMM methods suggests a positive effect.

We now turn to the effects of financial liberalization.[54] Figure 4.2 shows a positive relationship between the degree of financial liberalization and the relative size of the small enterprise sector. Table 4.4 shows an overall positive effect of capital account openness on the relative size of the small enterprise sector in a country. Both OLS levels and Within-Group estimates show that capital account openness has a positive effect on the relative size of the small enterprise sector although the relationship is significant only in the case of the Within-Groups estimate.

In column (3), we apply the differenced GMM estimation method with the full instrument set and the results indicate a positive and significant effect. However, the coefficient of the autoregressive parameter lies below the corresponding value obtained from the Within-Groups estimator suggesting that our estimate might be biased downward. In addition, the p-values for the test of first-order and second-order serial correlation suggest that the null hypotheses of no autocorrelation are not rejected in both cases. To address this issue, in column (4), we again apply the differenced GMM estimator but this time with a smaller set of instruments – i.e. we do not use any lagged levels of the series (y_{it}, x_{it}) or $(SEMSELE_{it}, KAOPEN_{it})$ further back than $t-4$ as instruments. The results, nevertheless, are still similar to the previous case.

In column (5), the system GMM estimation method is applied with a restricted set of instruments as in column (4). In other words, only lagged first-differences in the levels equation and lagged levels in the first-differenced equation at dates $t-2$, $t-3$ and $t-$

[54] We examined these effects using both the Chinn and Ito index of capital account openness and the Abiad and Mody index of financial liberalization. The results for the Abiad and Mody index indicate a positive but insignificant effect of financial liberalization on the relative size of the small enterprise sector. However, because data for only a small number of countries are available, this estimate might be subjected to small sample bias. For this reason, only the results from the estimation using the Chinn and Ito measure are discussed.

4 are used as instruments. This time the p-values for both the tests of autocorrelation suggest that there is first-order serial correlation but not second-order serial correlation, as expected. In addition, the Sargan test indicates that there is no problem due to invalid over-identifying restrictions. However, there is an issue related to the coefficient on the lagged dependent variable which lies above the corresponding value obtained from the OLS estimation method. This fact shows that our estimate is likely to be upward-biased. It might be due to the presence of serious finite sample biases.

To further examine the relationship using system GMM, we experiment with two more versions of system GMM using two different moment conditions. In column (6) only the lagged first-difference of SEMSELE is used as an instrument, while both the lagged first-difference and lagged level of KAOPEN are used. On the contrary, in column (7) both the lagged first-difference and the lagged level of SEMSELE but only the lagged first-difference of KAOPEN are used as instruments. This approach can help avoid overfitting and also reflects the possibility that the system GMM assumptions may be incorrect. The results from these experiments indicate that there are some improvements in the values of the autoregressive parameter. In these cases, the first-order and second-order serial correlation tests as well as the Sargan tests are also supportive.

Finally, the bias corrected least square dummy variable estimation method is applied and the results are presented in column (8). Similar to the system GMM estimation method, the LSDVC estimation method indicates that there is a positive and significant effect of capital account liberalization on the relative size of the small enterprise sector in the economy.

In general, the table provides evidence that financial liberalization in terms of capital account openness exerts positive and significant effects on the relative size of small enterprise sectors in the economy.

4.5 Sensitivity Check

In this section we check the robustness of our results by re-estimating our models on a restricted sample by dropping the first two three-year and the last three-year periods from our analysis.

Table 4.5 and Table 4.6 present results for our analysis of the effects of financial development, and of financial liberalization in terms of capital account openness, on the relative size of the small enterprise sector, respectively. Again, for comparison purposes we present results from OLS levels, the Within-Group, Differenced-GMM, System-GMM, and the bias-corrected Least Square Dummy Variable (LSDVC) estimates. Arrangements of the results are similar to those in Table 4.3 and Table 4.4.

Similar to our earlier findings, the results in Table 4.5 do not present any clear picture of the effect of financial development on the relative importance of small enterprises. The results from OLS levels, Within-Groups, differenced GMM estimation with a restricted set of instruments and LSDVC estimates in columns (1), (2), (4) and (8), respectively, suggest that financial development exerts a negative although weak effect on the relative size of the small enterprise sector. On the contrary, as shown in columns (3), (5), (6) and (7), estimating the relationship using differenced GMM estimation with a full set of instruments and system GMM estimation suggest a positive but insignificant effect.

With the exception of OLS levels estimation, where a negative sign on the coefficient is obtained, the results in Table 4.6 suggest positive effects of capital account openness on the relative size of the small enterprise sector. Nevertheless, these results are fragile, with only Within-Groups, differenced GMM estimation with a full set of instruments and LSDVC estimates yielding a statistically significant relationship. These findings can be expected because the use of a smaller sample for estimation of the models would result in less information: dropping time periods from our estimates is likely to cause standard errors to increase, resulting in less precise estimates. Nevertheless, the sensitivity of the results to dropping time periods means that we should be very careful in interpreting these estimation results.

Therefore, our model re-estimations using a restricted sample do not yield a clear picture of the impact of either the degree of financial liberalization or the level of domestic financial development on the relative size of the small enterprise sector.

4.6 Conclusion

This study examines the disproportionate effects of financial development and financial openness on small and large enterprises. This research contributes to the ongoing debate about whether financial development and financial liberalization disproportionately benefit small enterprises. Our estimation results suggest a positive impact of financial liberalization, as indicated by the degree of capital account openness, on the relative share of the small enterprise sector. Nevertheless, the strength of this relationship varies depending on the number of time periods used. When a larger sample is used, the results indicate that financial liberalization enhances the growth rate of small firm sectors more than large firm sectors. Nevertheless, results become fragile when a more restricted set of sample is used for estimation. Moreover, evidence for the effects of financial development is less clear.

Several existing papers suggest that financial liberalization contributes to economic growth by introducing greater efficiency and competition in credit markets. Such competition results in lower lending rates as well as an increase in the availability of credit to firms. Our findings, though fragile, support this view by showing a positive effect of capital account openness on the relative share of the small enterprise sector. In addition, our study supports the hypothesis that financial liberalization primarily boosts the growth of small firms more than large firms. This may be because generally large firms are the main benefactors under a repressed financial system through such policy arrangements as credit controls and credit rationing. Under these arrangements, credit may often go to larger and more politically connected firms at the expense of smaller firms. For these reasons, removal of financial repression policies would eliminate such distortions in credit allocation and, thus, should benefit smaller firms more than their larger counterparts.

A policy implication of these findings is that policies that enhance the degree of openness of the financial system could promote economic growth by boosting the growth rate of small enterprises in particular.

With regards to the contribution of domestic financial sector development to firm growth, some theories imply that financial development is especially beneficial to small firms by reducing transaction and information costs in the financial markets that hinder small firm growth. Other theories suggest that financial development is particularly beneficial for large firms. Unfortunately, our findings support neither view. In particular, we do not find any significant effect, positive or negative, of domestic financial development on the relative size of small enterprise sectors. A potential explanation for this finding is that financial development may provide similar benefits to the small and the large enterprise sectors, rather than favoring one sector over the other.

Table 4.1: The Financial Development Indicator (FINDEV)

3 Principal Components Retained

Component	Eigenvalue	Difference	Proportion	Cumulative
1	2.13296	1.52297	0.7110	0.7110
2	0.60999	0.35294	0.2033	0.9143
3	0.25705	-	0.0857	1.0000

Eigenvectors

Variable	1	2	3
BANK	0.51300	0.84039	0.17484
LLY	0.59175	-0.49379	0.63718
PRIVATE	0.62182	-0.22341	-0.75062

Notes: The table shows how our Financial Development indicator (FINDEV) is constructed based on the data from the three different measures of financial development, i.e. BANK, LLY and PRIVATE. FINDEV is constructed using the standardized first principal component of three standardized financial measures. This is the linear combination of the variables that has the highest sample variance, subject to the constraint that the sum-of-squares of the coefficients equals unity.

Table 4.2: List of Countries (61 countries)

DZA Algeria	MAR Morocco
ARG Argentina	NLD Netherlands
AUS Australia	NZL New Zealand
AUT Austria	NIC Nicaragua
BGD Bangladesh	NGA Nigeria
BEL Belgium	NOR Norway
BOL Bolivia	PAK Pakistan
BRA Brazil	PAN Panama
BGR Bulgaria	PER Peru
CMR Cameroon	PHL Philippine
CAN Canada	POL Poland
CHL Chile	PRT Portugal
COL Colombia	SGP Singapore
CRI Costa Rica	ESP Spain
CYP Cyprus	LKA Sri Lanka
DNK Denmark	SWE Sweden
ECU Ecuador	THA Thailand
EGY Egypt	TUR Turkey
SLV El Salvador	GBR United Kingdom
FIN Finland	VEN Venezuela
FRA France	ZWE Zimbabwe
GHA Ghana	
GRC Greece	
GTM Guatemala	
HUN Hungary	
ISL Iceland	
IND India	
IDN Indonesia	
IRN Iran	
IRL Ireland	
ISR Israel	
ITA Italy	
JPN Japan	
JOR Jordan	
KEN Kenya	
KOR Korea	
KWT Kuwait	
MYS Malaysia	
MLT Malta	
MUS Mauritius	

Notes: This table lists names and country codes of the sample countries used in this study. There are 61 countries including both developed and developing countries where sufficient data are available.

Table 4.3: Financial Development (FINDEV) and the Small Enterprise Sector (Whole Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: SEMPSELE	OLS	Within Group	DIFF-GMM Full	DIFF-GMM Reduced	SYS-GMM Reduced	SYS-GMM Reduced (SEMPSELE, Diff-Eq)	SYS-GMM Reduced (FINDEV, Diff-Eq)	LSDVC
SEMPSELE $_{t-1}$	0.950***	0.623***	0.408***	0.400*	0.882***	0.726***	0.996***	0.750
	(54.910)	(13.030)	(3.490)	(1.910)	(13.630)	(5.350)	(7.760)	(15.810)
FINDEV $_{t-1}$	-0.265*	-0.270	2.109**	2.199**	0.009	0.304	1.982*	-0.341
	(-1.890)	(-0.670)	(2.510)	(2.550)	(0.030)	(0.610)	(1.720)	(-0.600)
p-value for m1			0.19	0.32	0.03	0.06	0.04	
p-value for m2			0.69	0.64	0.70	0.79	0.86	
Sargan test (p-value)			0.91	0.66	0.76	0.71	0.61	
No. Obs.	357	357	303	303	357	357	357	303
No. Group		51	51	51	54	54	54	51

Notes: numbers in parenthesis are t-statistics, except in the case of LSDVC where the z-statistics are shown. m1 and m2 are tests for first-order and second order serial correlation. The Sargan test is a test of the validity of the over-identifying restrictions for the GMM estimators. The GMM results are two-step estimates with heteroskedasticity-consistent standard errors and test statistics. "Full" indicates that all available lags are used as instruments. "Reduced" suggests that a restricted set of instruments is applied; and in this case no lag of the dependent variable further back than t-4 is used. "SEMSELE, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of SEMSELE is used as an instrument while both the lagged first-difference and the lagged level of FINDEV are used. "FINDEV, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of FINDEV is used as an instrument while both the lagged first-difference and the lagged level of SEMSELE are used.

Table 4.4: Capital Account Openness (KAOPEN) and the Small Enterprise Sector (Whole Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: SEMPSELE	OLS	Within-Group	DIFF-GMM Full	DIFF-GMM Reduced	SYS-GMM Reduced	SYS-GMM Reduced (SEMPSELE, Diff-Eq)	SYS-GMM Reduced (KAOPEN, Diff-Eq)	LSDVC
SEMPSELE $_{t-1}$	0.954 (67.850)	0.656 (16.000)	0.504 (3.260) ***	0.144 (0.620)	1.003*** (10.840)	0.879*** (9.380)	0.964** (6.560)	0.796*** (15.270)
KAOPEN $_{t-1}$	0.025 (0.230)	0.860*** (3.490)	2.207** (2.390)	2.474** (2.370)	0.599** (2.160)	0.382 (1.090)	1.691** (2.010)	0.843*** (2.740)
p-value for m1			0.11	0.85	0.01	0.02	0.02	
p-value for m2			0.83	0.45	0.61	0.61	0.78	
Sargan test (p-value)			0.56	0.43	0.36	0.29	0.42	
No. Obs.	424	424	365	365	424	424	424	365
No. Group		59	59	59	59	59	59	59

Notes: numbers in parenthesis are t-statistics, except in the case of LSDVC where the z-statistics are shown. m1 and m2 are tests for first-order and second order serial correlation. The Sargan test is a test of the validity of the over-identifying restrictions for the GMM estimators. The GMM results are two-step estimates with heteroskedasticity-consistent standard errors and test statistics. "Full" indicates that all available lags are used as instruments. "Reduced" suggests that a restricted set of instruments is applied; and in this case no lag of the dependent variable further back than t-4 is used. "SEMSELE, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of SEMSELE is used as an instrument while both the lagged first-difference and the lagged level of KAOPEN are used. "KAOPEN, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of KAOPEN is used as an instrument while both the lagged first-difference and the lagged level of SEMSELE are used.

Table 4.5: Financial Development (FINDEV) and the Small Enterprise Sector (Sub-Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: SEMPSELE	OLS	Within Group	DIFF-GMM Full	DIFF-GMM Reduced	SYS-GMM Reduced	SYS-GMM Reduced (SEMPSELE, Diff-Eq)	SYS-GMM Reduced (FINDEV, Diff-Eq)	LSDVC
SEMPSELE $_{t-1}$	0.957*** (51.38)	0.549*** (9.510)	0.461*** (3.960)	0.441** (2.390)	0.864*** (11.640)	0.744*** (5.620)	0.906*** (7.870)	0.762*** (10.630)
FINDEV $_{t-1}$	-0.250 (-1.680)	-0.158 (-0.310)	0.971 (0.700)	-0.425 (-0.380)	0.117 (0.370)	0.326 (0.650)	0.839 (0.970)	-0.523 (-0.780)
p-value for m1			0.08	0.14	0.05	0.06	0.05	
p-value for m2			0.88	0.72	0.77	0.81	0.84	
Sargan test (p-value)			0.31	0.54	0.63	0.61	0.50	
No. Obs.	273	273	260	260	273	273	273	179
No. Group		53	49	49	53	53	53	49

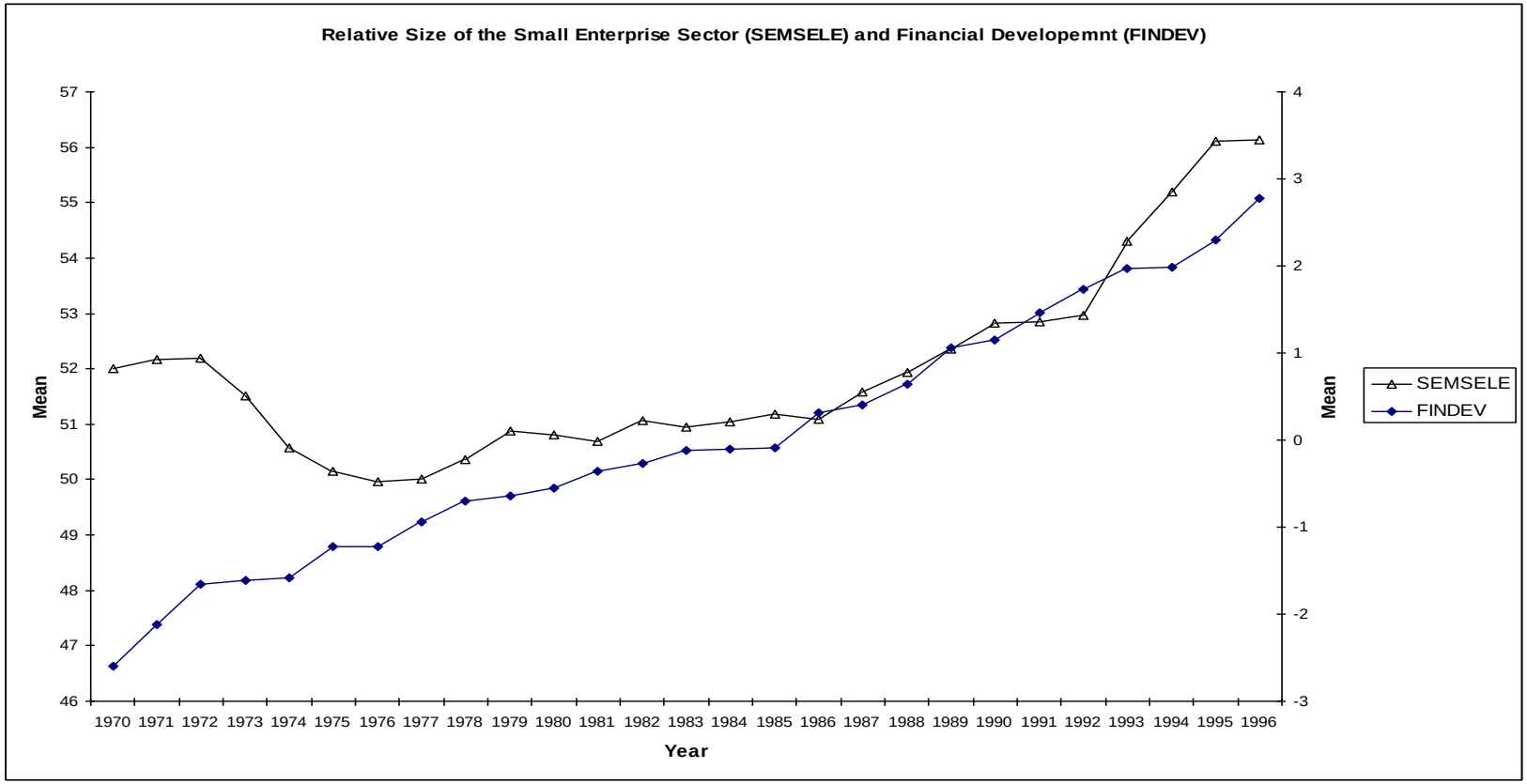
Notes: numbers in parenthesis are t-statistics, except in the case of LSDVC where the z-statistics are shown. m1 and m2 are tests for first-order and second order serial correlation. The Sargan test is a test of the validity of the over-identifying restrictions for the GMM estimators. The GMM results are two-step estimates with heteroskedasticity-consistent standard errors and test statistics. "Full" indicates that all available lags are used as instruments. "Reduced" suggests that a restricted set of instruments is applied; and in this case no lag of the dependent variable further back than t-4 is used. "SEMSELE, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of SEMSELE is used as an instrument while both the lagged first-difference and the lagged level of FINDEV are used. "FINDEV, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of FINDEV is used as an instrument while both the lagged first-difference and the lagged level of SEMSELE are used.

Table 4.6: Capital Account Openness (KAOPEN) and the Small Enterprise Sector (Sub-Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: SEMPSELE	OLS	Within-Group	DIFF-GMM Full	DIFF-GMM Reduced	SYS-GMM Reduced	SYS-GMM Reduced (SEMPSELE, Diff-Eq)	SYS-GMM Reduced (KAOPEN, Diff-Eq)	LSDVC
SEMPSELE $_{t-1}$	0.959*** (64.300)	0.614*** (11.900)	0.496*** (3.200)	0.293 (1.130)	0.947*** (12.980)	0.875*** (10.650)	0.907*** (6.400)	0.819*** (11.620)
KAOPEN $_{t-1}$	-0.005 (-0.040)	0.735** (2.570)	1.464** (2.190)	1.499 (1.220)	0.444 (1.340)	0.132 (0.310)	1.122 (1.370)	0.580* (1.850)
p-value for m1			0.07	0.31	0.02	0.02	0.03	
p-value for m2			0.78	0.79	0.64	0.61	0.75	
Sargan test (p-value)			0.61	0.62	0.43	0.31	0.38	
No. Obs.	325	325	316	316	325	325	325	212
No. Group		59	59	59	59	59	59	58

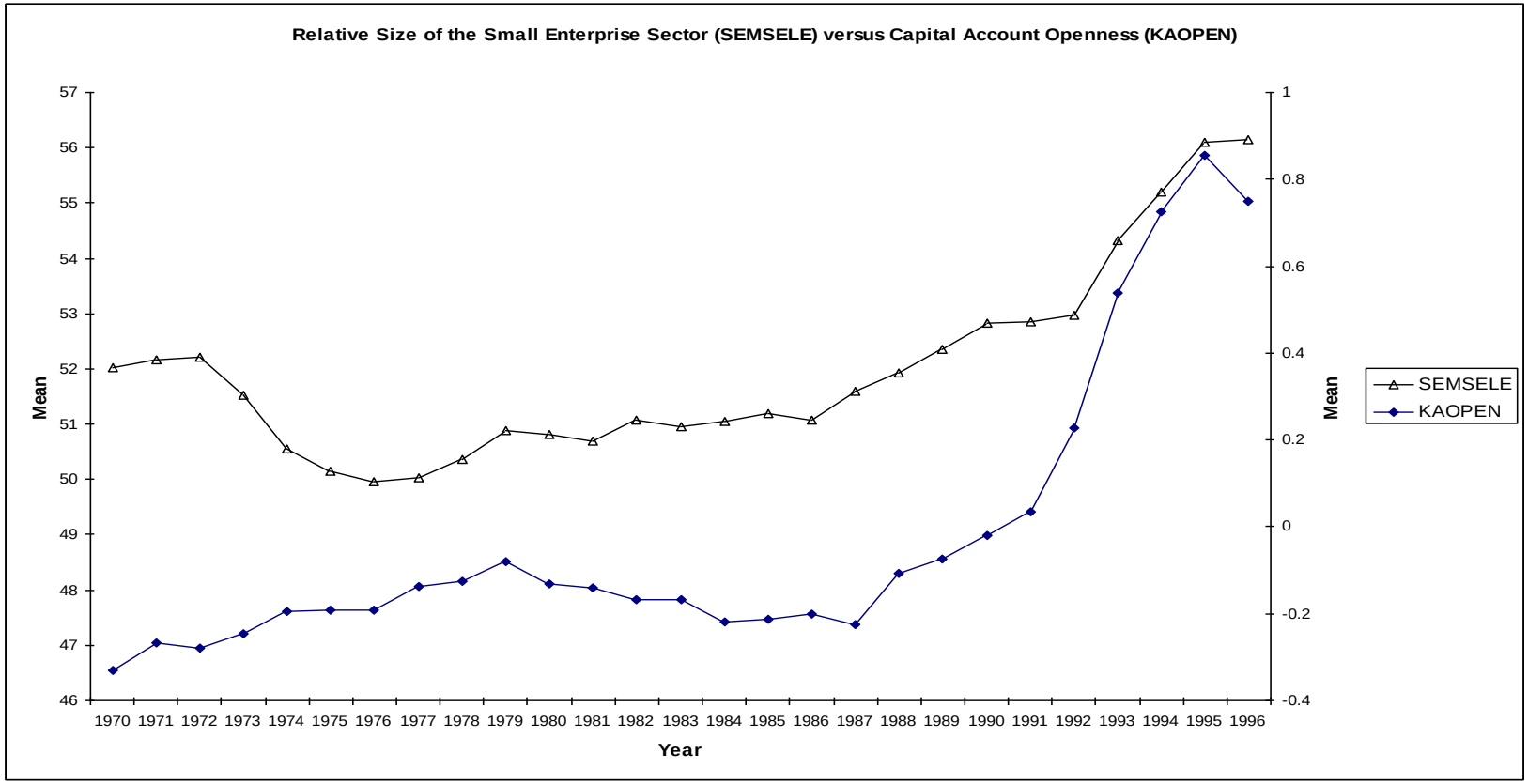
Notes: numbers in parenthesis are t-statistics, except in the case of LSDVC where the z-statistics are shown. m1 and m2 are tests for first-order and second order serial correlation. The Sargan test is a test of the validity of the over-identifying restrictions for the GMM estimators. The GMM results are two-step estimates with heteroskedasticity-consistent standard errors and test statistics. "Full" indicates that all available lags are used as instruments. "Reduced" suggests that a restricted set of instruments is applied; and in this case no lag of the dependent variable further back than t-4 is used. "SEMSELE, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of SEMSELE is used as an instrument while both the lagged first-difference and the lagged level of KAOPEN are used. "KAOPEN, Diff-Eq" indicates that the following moment conditions are applied: only the lagged first-difference of KAOPEN is used as an instrument while both the lagged first-difference and the lagged level of SEMSELE are used.

Figure 4.1: Trends of the Relative Size of the Small Enterprise Sector and of Financial Development



Notes: This figure shows the changes over time in the relative size of the small enterprise sector and of the level of financial development across countries. The values on the left y-axis are the cross-country averages of the relative size of the small enterprise sector, while the right y-axis indicates the average values of the aggregate financial development index. Years are denoted on the x-axis.

Figure 4.2: Trends of the Relative Size of the Small Enterprise Sector and of Capital Openness (KAOPEN)



Notes: This figure shows the changes over time in the relative size of the small enterprise sector and of the degree of financial liberalization (openness) across countries. The values on the left y-axis are the cross-country averages of the relative size of the small enterprise sector, while the right y-axis indicates the average values of the aggregate financial development index. Years are denoted on the x-axis.

Chapter 5
The Effects of Institutions on the Firm Size Distribution

A limited number of studies have been conducted to examine the role of institutions in explaining cross-industry and cross-country variation in the size distribution of firms. Studies including that of Davis and Henrekson (1997) have looked into the relationship between the regulatory environment in a country and the firm size distribution. Other studies take a slightly different approach by examining the indirect effects of institutions on firm size structure.[55] For example, they look at the potential roles of institutions in facilitating the development of the financial system, which, in turn, may affect the growth of firms of different sizes. While these studies provide some informative answers about the association between institutions and the firm size distribution, they fall short of providing a comprehensive analysis regarding the potential causality of this relationship.

This study, thus, aims to address this issue. Specifically, we will examine cross-country evidence to see whether a causal relationship between institutional quality and the share of the small and medium enterprise (SME) sector exists. Furthermore, we will determine potential disproportionate impacts of institutions on firms of different sizes by analyzing the relationship between a measure of institutional quality and the relative size of the small enterprise sector versus the large enterprise sector.

Instrumental variable (IV) estimation methods are employed in our analysis. The results indicate that although our measure of institutions has a positive relationship with the share of the SME sector, this relationship is not robust to controlling for simultaneity bias. This suggests that while countries with good institutions generally have a large SME sector, there is insufficient evidence to support the view that institutional quality exerts a causal effect on the size of the SME sector. Furthermore, when the UNIDO-based measure of the relative size of the small enterprise sector is used as the dependent variable, we do not find any evidence that indicates potential disproportionate effects of institutions on small and large sized firms.

---

[55] See, for instance, Beck et al. (2000).

This paper contributes to research on the institutional relationship with firm structure in a number of ways. The use of instrumental variable estimation for the analysis enables us not only to determine the correlation between institutions and the SME share but also to capture any potential causal effect that our institutional measure may have on the importance of the small and medium enterprises in the economy. Moreover, as evidence suggests that small and medium enterprises play important roles in the economic growth and development process, the findings from this study may have implications for policy makers by enabling them to design and implement appropriate policy measures which promote the growth of this sector.

The remainder of this chapter is organized as follows. The next section is the literature review where we will discuss some of the previous studies of the relationship between institutions and the firm size distribution. Section 5.2 will describe the sample of countries used in this study as well as the different variables and their respective sources. Next, we proceed to discuss the methodologies and the results from our analysis of the relationship between our measure of institutions and the share of the SME sector in Section 5.3. We then present the results from regression estimation of the relationship between institutions and UNIDO-based measure of the relative size of the small enterprise sector versus the large enterprise sector in Section 5.4. Finally, Section 5.5 is the conclusion.

5.1 Previous Studies

Political institutions can affect the firm size distribution through various channels. The application of particular regulations can tilt the playing field towards small firms and away from large firms, or vice versa. Davis and Henrekson (1997), for instance, examine the effects of Swedish institutional structure on firm size in terms of the distribution of employment across sectors.[56] Looking particularly at the role of the economic policy environment in Sweden – as determined by high business taxes, employment security laws, highly regulated credit markets, the mandatory national pension system, central wage-setting institutions associated with highly compressed wages, and the rapidly increasing size of the

[56] They look at institutional structure in Sweden relative to other European countries and the United States.

public sector – they find that the overall policy environment in Sweden strongly disfavored less capital intensive, smaller firms. Such findings give a good explanation why the Swedish economy is dominated by large firms.[57]

Institutional efficiency and effectiveness can work to relax constraints on firms by leveling the playing field for firms of all sizes. This is consistent with the finding of Kumar, Rajan and Zingales (1999). They show that countries with better institutions, as measured by judicial system efficiency, tend to have lower dispersion in firm size within an industry. They also find a positive correlation between the quality of institutions and average firm size. Likewise, Beck, Demirguc-Kunt and Maksimovic (2003) argue that there is a significantly positive relationship between firm size and a country's legal system. They find that firms are larger in countries with more efficient legal systems. However, their study is slightly different from this chapter in the sense that their main focus is on cross-country difference in firms' absolute size and not in the share of small and medium enterprises in the manufacturing sector.

Beck, Demirguc-Kunt and Levine (2002) find that financial, legal and corruption constraints negatively affect firm growth. However, the extent to which these factors constrain firm growth depends very much on the size of firm and it is consistently the smallest firms that are most adversely affected by all three constraints. Consequently, institutional development that helps improve the financial system or the legal system or reduce corruption would provide the most benefit for the most-constrained groups of small and medium size firms.

Additionally, political and regulatory institutions can have an effect on firm growth and on the firm size distribution via their role in fostering the development of the financial system. Such financial system development, in turn, would relax financial constraints on firms, especially on smaller firms, and allow them to grow. Demirguc-Kunt and Maksimovic (1999) find that the efficiency of a legal system facilitates firms' access to external finance needed for growth. King and Levine (1993), Levine and Zervos (1998) and Beck,

[57] On a similar note, Ringleb and Wiggins (1990) study the effects of changes in liability laws in the United States on firm structure. They find the effect to be important in explaining the time variation of size of firms. They find evidence that the introduction of liabilities laws causes firms to seek ways to protect their assets by creating separate legal entities that can be put under the protection of limited liability laws. This has lead to a large increase in the number of small corporations in hazardous sectors.

Levine and Loayza (2000) show that financial development promotes growth and that different legal origins explain differences in financial development. In the same way, La Porta et al. (1997) show that countries with poorer investor protection, measured by the character of legal rules and the quality of law enforcement, have smaller and narrower capital markets which are not conducive to firm growth. Our study differs from these earlier studies in the sense that we seek to investigate the direct relationship between institutional quality and the size of the SME sector instead of an indirect relationship through financial sector development, by controlling for the level of financial sector development within each country. Finally, unlike some previous studies which look at a country-specific regulatory environment, we use evidence from a number of countries to determine the potential effects of institutions on the SME sector size.

5.2 Sample Countries and Data

This section discusses different variables used in this study as well as their respective sources.

5.2.1 Dependent Variable

For our dependent variable, we use the measure of the relative importance of small and medium enterprises constructed by Ayyagari, Beck and Demirguc-Kunt (2003) based on the share of manufacturing employment accounted for by small and medium enterprises in the economy, SMEOFF. The data for SMEOFF are from 76 countries and are averaged over the 1990-1999 period. We also use as the dependent variable the UNIDO-based measure of the relative size of the small enterprise sector versus the large enterprise sector, SEMSELE, to capture potential disproportionate effects of institutions on small enterprise sectors versus large enterprise sectors. The construction of this measure was described in chapter 3. The data are for a sample of 57 countries and are averaged over the period 1990-1996. Table 5.1a and b provide lists of countries used in the analysis with SMEOFF and SEMSELE, respectively.

5.2.2 Measures of Institutions

A number of quantitative measures of political institutions have been constructed and used in various empirical studies. One such measure is the Polity IV index put together by Marshall and Jaggers (2002). This index determines different political regime characteristics by measuring the constraints placed on the executive power. Another institutional data set is the survey indicators of institutional quality from the *International Country Risk Guide (ICRG),* which is produced by a private company for sale mainly to firms and portfolio managers who are considering foreign investments (Jalilian, Kirkpatrick and Parker, 2003). This dataset has also been used in many studies including Knack and Keefer (1995), Hall and Jones (1999) and Neumayer (2002). These data are produced annually and include risk assessments for international investors along such dimensions as law and order, bureaucratic quality, corruption, risk of expropriation by the government, and risk of government contract repudiation. Each variable is measured on a point scale with higher points denoting better performance with respect to the variable concerned. The assessment is based on expert analysis from an international network and is subject to peer review.

The third set of institutional data is an index measure of government effectiveness compiled initially by Kaufmann, Kraay and Zoido-Lobaton (1999) and updated by Kaufmann, Kraay and Mastruzzi (2005). This set of data consists of index measures for six different dimensions of governance including voice and accountability, political stability and violence, government effectiveness, regulatory quality, rule of law and control of corruption.[58] Voice and Accountability measures political, civil and human rights within a country. Political Stability and Violence indicates the likelihood that violent threats or changes in government, including terrorism, are likely to take place within a country. Government Effectiveness is used to assess the degree of competence of the bureaucracy and the quality of public service delivery, while Regulatory Quality indicates the incidence of market-unfriendly policies. Another governance indicator, Rule of Law, gauges the quality of contract enforcement, the police, and the courts, as well as the likelihood of crimes and violence. Finally, Control of Corruption measures the exercise of public power for private

[58] The construction of these six governance dimensions are based on a total of 352 individual variables measuring the perceptions of governance, drawn from 37 separate data sources constructed by 31 different organizations. See Kaufmann, Kraay and Mastruzzi (2005).

gain, including both petty and grand corruption and state capture. These indexes are available for every other year from 1996 to 2004. Each of these index variables range from -2.5 to 2.5, with higher numbers corresponding to better governance and higher institutional quality.

Out of the three sets of data for institutions, the Polity IV index may be the least suitable for our analysis since it only indicates the restraints put on government but not the effectiveness or efficiency of government institutions in promoting firm growth. For example, a country such as the Philippines may have a higher degree of restraint on executive power than a more authoritarian state like Singapore. However, it has less effective government institutions that are more prone to corruption than Singapore, which creates a tougher environment for firms to grow.[59]

The remaining two sets of data are similar because they both take into account various factors affecting the efficiency and effectiveness of government institutions as well as different aspects of political and regulatory environments within a country. In this study we focus on the measures of government effectiveness from the Kaufmann et al. dataset because it takes into account a larger number of factors that may affect governance and is thus more suitable than the ICRG data for our purpose.

We construct an aggregate index measure of institutional quality as the average of the six Kaufmann et al. measures of governance for the year 1996. This new aggregate variable is labeled as INST, indicating the quality of institutions for each country in 1996.[60] By default, this new variable ranges between -2.5 and 2.5 where higher numbers correspond to better institutional quality. The correlations between our aggregate index of institutional quality measure and each individual measure of governance are displayed in Table 5.2.

---

[59] A good example would be to look at a specific period, say between 1993 and 2003, when the Philippines scored consistenly higher in terms of the amount of contraints placed on the executive power with a score of 6 compared to Singapore with a score of only 3. The scores are based on the scale from 1 to 10, with 10 indicating the most constraints. However, evidence suggests that during that same period, Singapore fared much better than the Philippines in terms of the effectiveness of governement institutions, the ability to deal with corruption and overall economic performance.

[60] The year 1996 is chosen because it is situated in the middle of the period 1990-99 examined in this study. We also experiment with the data for 1998, which is the other year where data is available and which falls in our study period. We obtain similar results.

5.2.3 Independent Variables

In addition to our measure of institutional quality, we also control for a number of economic, social and geographical factors. Specifically, the list of control variables includes the log of income per capita (LGDP) and its quadratic form ($LGDP^2$); average years of schooling (SCH); log of exports share in GDP (LEXPGDP); continental dummy variables for Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM); service sector share in GDP (SERVGDP); closest air distance to a major port (AIRDIST); civil rights index which assesses the degree of protection of vulnerable groups against employment discrimination (CIVILR); employment laws index of the protection of labor and employment laws (EMPLAW) and the degree of financial system development in a country (FINDEV).

For those variables with missing values, we fill in the missing values by using a regression-based imputation method so as to maximize the number of observations for our econometric analysis of the dataset. The idea is to generate the missing value for a variable based on the values of other variables which are present using a linear regression. However, we realize the limitation of this method in providing estimation of the missing values and the potential impact that errors may have on the actual econometric analysis, especially for those variables with a fairly large number of missing values. Thus, we only impute those variables with less than ten missing values in order to minimize the impact that these estimated values may have on our analysis. Those variables of interest with more than ten missing values are dropped.

Table 5.3 provides a summary description of each variable as well as their respective sources. Table 5.4 provides descriptive statistics of all the variables used in our analysis. It shows that there is a large variation in the employment share of the SME sector across the countries in our sample, ranging from 4.59 in Belarus to 86.70 in Thailand. The cross-country variation in the quality of institutions is also fairly wide ranging from -1.75 in Tajikistan to 1.93 in New Zealand.

Table 5.5 shows the correlation between the dependent variable, the institutional measures and other variables used for this analysis. Simple correlations indicate a positive re-

lationship between institutional quality and the share of SME sector with a correlation of 0.41.

5.3 Methodology and Results: INST versus SMEOFF

5.3.1 Ordinary Least Squares Estimation

To determine the relationship between our institutional variable and the size of the small and medium enterprise sector, we run a simple Ordinary Least Squares (OLS) regression of the size of the small and medium enterprise sector on the measure of institutional development and a number of conditioning variables. The linear regression takes the form:

$$SMEOFF_i = \alpha + \beta\, INST_i + \gamma\, X_i + u$$

where $SMEOFF$ is the measure of the size of the small and medium enterprise sector, $INST$ is the institutional quality indicator, X is a set of conditioning variables which relate to the size of the small and medium enterprise sector, i is the country index and u is the error term.

Table 5.6 displays the results of the Ordinary Least Squares regression. Column 1 shows that there is a positive and significant relationship between the share of small and medium enterprises and our measures of institutional quality. Moreover, the size of the coefficient also suggests a strong and economically meaningful relationship. For example, an increase (decrease) of 1 standard deviation in the quality of institution is associated with an increase (decrease) of more than $\frac{1}{2}$ of a standard deviation in the relative importance of the small and medium enterprise sector in the economy, which is fairly significant. Figure 5.1 shows this relationship for our sample of 76 countries and it confirms a positive correlation between institutional quality and the share of the SME sector. The R^2 value suggests that institutional quality accounts for about one-third of the cross-country variation in the share of small and medium enterprises in manufacturing sector employment.

We include the log of income per capita in 1996 (LGDP) as a proxy for the level of economic development to control for the risk that the institutional variables being used are proxying for other factors that depend on a country's level of development. Its quadratic form, $LGDP^2$, is also included to examine if there exists a non-linear relationship between share of small and medium enterprise sector and income level. The income data are obtained from the Penn World Table Mark 6.1 compiled by Heston et al. (2002).[61] Column 2 shows regression results when both log of income per capita and its quadratic form are added. The inclusion of these two variables increases the coefficient of the institutional measure while maintaining statistically significant relationship with the dependent variable. Interestingly, the income measure is found to have a positive but insignificant correlation with the dependent variable, while its quadratic form has negative but insignificant relationship.

In Column 3, we include measures of human capital and a country's exposure to international markets. In particular, the average years of schooling for the population aged 15 and over (SCH) are used to capture the effect of the level of human capital accumulation and are obtained from Barro and Lee's educational attainment dataset.[62] A higher level of human capital in a country may either promote larger firms, due to higher managerial skills, or more and thus smaller firms, due to widely available entrepreneurial skills (Rosen, 1982; Kremer, 1993). As for a country's exposure to international market, we use log of export share in GDP (LEXPGDP) as an indicator of a country's trade with outside world. The inclusion of these additional controls does not have much effect on the relationship between institutional quality measure and income level and the dependent variable. On the other hand, human capital is found to have a negative and significant relationship with the size of the small and medium enterprise sector. Moreover, our measure of openness registers a negative and significant coefficient suggesting that more open countries tend to have larger firms.

---

[61] There is, however, no consensus regarding the direct relationship between firm size and the level of economic development. For instant, conventional wisdom suggests that richer countries should have larger frims because potential entrepreneurs face higher opportunity costs in the form of higher wages (Lucas, 1978). On the contrary, studies find little evidence that support such a claim (Kumar et al., 1999).

[62] Average years of schooling in 1995 is used instead of 1996 (the year selected for INST) because of the availability of data.

In Column 4, we add dummy variables for countries in Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM) in order to control for a possible association between geographical characteristics and the firm size distribution. Column 5 includes service sector share in GDP (SERVGDP) which is used as an indicator of the degree of de-industrialization that takes place within the economy; the closest air distance to a major port (AIRDIST) representing a country's relative geographical remoteness; civil rights index which assesses the degree of protection of vulnerable groups against employment discrimination (CIVILR) and employment laws index of the protection of labor and employment laws (EMPLAW). Again, results in Columns 4 and 5 show positive and statistically significant relationship between institutional quality measure and the share of small and medium enterprise sector in the economy. Similarly, relationship between log of export share in GDP and the dependent variable remains negative and statistically significant. Furthermore, addition of these new variables result in both log of income per capita and its quadratic form having statistically significant relationships with the SME sector share. However, the signs on the coefficients of both terms are opposite with the log of income and its squared value having positive and negative signs, respectively. The opposite signs may suggest that there exist a non-linear, inverted U-shape relationship between the SME sector share and income level; and that there may be a decline in the relative importance of small and medium enterprise sector in the economy in the long run as the country develops.

Finally, since we are interested in examining a direct relationship between institutional quality and firm size, we include a measure of financial development (FINDEV) to account for the potential indirect effects that institutions might have on the firm size distribution through an improvement in the financial system.[63] Column 6 shows the results when FINDEV is added to the regression. Columns 6 shows that inclusion of this additional control does not have much effect on our main estimates. Figure 5.2 illustrates a partial relationship between the measure of institutional quality and the share of the SME sector for our sample of 76 countries. This figure shows the relationship to be positive after conditioning on various controls mentioned above.

[63] See Chapter 2 for a more detailed explanation of how FINDEV is derived.

Additionally, we present the results of the OLS estimation of the relationship between each component of the institutional measure and the SME share. Panel A of Table 5.7 uses the full set of conditioning variables while the log of income per capita and its quadratic form are dropped from the set of conditioning variables in Panel B. The results show that four of the six variables register a positive and significant relationship with the dependent variable. Political stability and the control of corruption index have positive but insignificant relationships with the dependent variable. Figure 5.3 shows the corresponding partial correlations.

5.3.2 Sensitivity Check with Extreme Bound Analysis

To further test the robustness of the relationship between the size of the small and medium enterprise sector and the measure of institutional quality, we follow Levine and Renelt (1992) in using the extreme-bound analysis (EBA) method proposed by Edward Leamer (1983).

The advantage of using this method is that it allows us to conduct a more rigorous robustness check of the relationship between our variables of interest by experimenting with a large number of possible combinations of our conditioning variables. Extreme-bound analysis involves first choosing and running a base regression which includes as explanatory variables the variable of interest and a number of other variables that are commonly used in previous studies. In this case, our variable of interest is our measure of institutional development (INST) and the common variables chosen are the log of income level (LGDP) and average years of schooling (SCH).[64] The next step is to change the subset of conditioning variables in a regression to determine the widest range of coefficient estimates on our variable of interest that is not rejected by standard hypothesis tests. If the coefficient of our variable of interest remains significant and of the same sign at the extreme bounds, then we can maintain with a fair amount of confidence that the result is robust.[65]

---

[64] Based on various prior growth research including those by Beck, Levine and Loayza (2000) and Beck, Demirguc-Kunt and Levine (2003), we choose the log of income level and our measure of human capital as the common variables to be included in the base regression.

[65] Please refer to the paper by Levine and Renelt (1992) for a more detailed description of the Extreme Bound Analysis (EBA) method.

To keep the regressions manageable we limit the number of variables in each subset of conditioning variables to four. Thus there are no more than seven independent variables per regression.

Panel A of Table 5.8 shows the results for regressions when extreme-bound analysis is applied. The coefficients of our measure of institution (INST) in the base, the extreme high bound and the extreme low bound regressions remain positive and robust. At the higher bound, the coefficient on INST is 15.482 with a t-statistic of 3.099. At the lower bound, the coefficient on INST is 9.423 with a t-statistic of 2.013. This suggests that the relationship between our measure of institutional quality and the share of small and medium enterprise sector is robust to changing sets of conditioning variables.

We also conduct further tests using an extreme bound analysis when the log of income and average years of schooling are not treated as common variables in the base regression, but are potentially included as conditioning variables. In other words, the base regression only contains our measure of institutional quality as the explanatory variable. In this case, we also restrict the number of variables in each subset of the conditioning variables to four, so that there will be five explanatory variables per regression, including our measure of institutions. The results are presented in Panel B. The results are similar to those shown in Panel A, with the coefficient on INST at the extreme high bound is 17.647 with a t-statistic of 4.512. At the extreme low bound, the coefficient on INST is 7.178 with a t-statistic of 2.073. In addition, the sign of the coefficient remains positive for both the extreme low bound and the extreme high bound regressions as well as for the base regression.

Therefore, the overall results from OLS and Extreme Bound Analysis do suggest a robust relationship between our measure of institutional quality and the share of small and medium enterprises in the manufacturing sector. Nevertheless, it is important not to interpret this strong correlation as causal. One plausible reason is that there are potentially several omitted determinants of the SME share that may be correlated with institutions. This, in turn, can introduce a positive bias in the OLS estimates. In addition, as Acemoglu et al. (2001) point out, the fact that the institutions variable is measured with considerable error and corresponds poorly to the cluster of institutions that matter in practice creates attenuation and may bias the OLS estimates downwards. All of these problems could be solved by using a good instrument for institutions.

5.3.3 Instrumental Variable Estimation

Our cross-section analysis may also be subjected to a simultaneity problem. To address this issue and to determine the effect of institutional development on the small and medium enterprise sector, we use a two-stage least squares instrumental variable (IV) estimation method to isolate the exogenous component of the institutional variable. A good instrument must play an important role in accounting for cross-country variation in the quality of institutions, but have no direct effect on the firm size distribution. In other words, the instrument must be uncorrelated with the error term.

There are a few studies which attempt to deal with the endogeneity of institutions through the use of instruments. Mauro (1995) uses an index of ethnolinguistic fragmentation as an instrument for corruption, arguing that ethnolinguistic fragmentation is highly correlated with corruption and other institutional variables, but can be assumed to be exogenous both to economic variables and to institutional efficiency.[66] However, the use of this instrumental variable has been criticized by Acemoglu et al. (2001) because, they argue, language diversity itself may be endogenous because such fragmentation almost completely disappeared in Europe during the era of growth when a centralized state and market emerged.[67]

Hall and Jones (1999) instrument for social infrastructure using absolute latitude (ABSLATIT) since, they argue, this geographical characteristic is correlated with the extent of Western influence, which may have led to good institutions.[68] In addition, empirical analysis conducted by Ayyagari et al. (2003) shows that there is no significant direct relationship between absolute latitude and the size of the SME sector, which supports the use of absolute latitude as a candidate instrument for the SME share.

Acemoglu, Johnson and Robinson (2001) suggest that the pattern of colonial settlement serves a very important role in understanding each country's political institutions because European settlers normally brought effective European institutions to the place

---

[66] The index of ethnolinguistic fractionalization measures the probability that two persons drawn at random from a country's population will not belong to the same ethnolinguistic group.

[67] See also Weber (1976) and Anderson (1983).

[68] By social infrastructure, they mean the institutions and governemnt policies that determine the economic environment within a country.

where they settled. On the other hand, where they did not settle, European colonizers instituted systems of arbitrary rule and expropriation of the local population. They further argue that the mortality of early European settlers in the countries they colonized played a key role in shaping their decision to settle or not. For this reason, Acemoglu et al. (2001) chose settler mortality as an instrument for modern day political institutions.

Along the same line of argument that European expansion influenced colonized countries' modern day institutions, La Porta et al. (1997, 1998, 1999) use legal origins as instruments for institutions. They argue that legal systems which were brought over by European colonizers into their subject countries have a strong influence on these countries' modern day legal structures and institutions.

Additionally, Putterman (2006) compiled estimates of the year when a country made a transition to agriculture based on various factors including archeological records on different geographical locations. We designate this variable as AGYEAR. Historical evidence suggests that the technological, social and economic development of the world's societies from Neolithic times to the present involves multiple steps beginning with a transition to a more densely settled agrarian state with currencies and taxation and finally to the development of modern enterprise systems, markets, and public sectors (Boserup, 1965; Diamond, 1998). Thus, we see that the transition to agriculture gives rise to a more complex social structure which requires the introduction of new ways to govern this more complex society and human interaction. This, in turn, would give rise to the creation of early institutional structures in these societies which, over time, evolve into modern day institutions.

In addition, it has been hypothesized that differences among human societies in the time at which the transition from reliance upon hunting and gathering to agriculture took place led to differences in levels of technological development and social organization that persisted into the era of European expansion beginning in the 15th century and into the present.[69] Therefore, the agricultural transition year could be a good indicator of a country's early development which has effects on its modern day institutions.

Because there is no perfect instrument for institutions, we will experiment with different instrumental variable sets comprising the instruments discussed above, except ethno-

[69] See, for example, Diamond (1998).

linguistic fragmentation and settler mortality.[70] We exclude ethnolinguistic fragmentation due to its potential endogenous nature as discussed by Acemoglu et al. (2001), while settler mortality is dropped because of the very limited number of observations available for analysis when settler mortality is included in our instrument set. Specifically, when settler mortality is included in the instrument set, the number of observations is reduced dramatically from 76 to 29. Such a limited number of observation might render any result uninformative. Therefore, the set of instruments used in this analysis will be drawn from data on the different legal origins (French, UK, German and Scandinavian legal origins), the agriculture transition year (AGYEAR) and absolute latitude (ABSLATIT).

In the output tables, we present the results from instrumental variable estimation as well as test statistics to test the validity and strength of each set of instrumental variables. To assess the validity of the instrumental variables, we provide p-values for the Sargan test of the overidentification restrictions. This test statistic allows us to determine whether the instrumental variables are associated with the share of small and medium enterprise sector beyond their role in explaining cross-country variation in the quality of institutions. This test is conducted under the null hypothesis that the excluded instruments (i.e. the instruments not included in the second stage regression) are valid and that the excluded instruments are correctly excluded from the estimated equation. A failure to reject the null hypothesis implies the failure to reject the validity of the instruments. This, in turn, would imply a failure to reject the view that the coefficient estimates capture the causal impact of institutional quality on the share of the SME sector.

We also provide in the tables first-stage F-statistics for all of the instruments. That is we test the null hypothesis that the instruments do not explain cross-country variation in the quality of institutions. This test statistic is thus used to determine the strength of our instruments. As various authors including Temple and Woßmann (2006) have pointed out, when instruments are only weakly correlated with the endogenous explanatory variables, instrumental variable estimation may be biased in small samples, and the conventional asymptotic approximations used for hypothesis tests and confidence intervals are likely to be unreliable.

---

[70] We use the *ivreg2* command in Stata to implement the 2SLS instrumental variable estimation method. See Baum, Schaffer and Stillman (2003).

Panel A of Table 5.9 shows the second-stage results from the instrumental variable estimation of the effects of institutional quality on the share of the small and medium enterprise sector. Panel B shows the corresponding first stages. We use five different sets of instruments. Column 1 uses different legal origins – UK legal origin, French legal origin, German legal origin and Scandinavian legal origin – as instruments; Column 2 uses legal origins and agricultural transition year; Column 3 uses legal origins and absolute latitude; Column 4 uses agricultural transition year and absolute latitude and, finally, Column 5 uses legal origins, agricultural transition year and absolute latitude as instruments.

In addition, conditioning variables used in both first-stage and second-stage regressions include log of income per capita (LGDP) and its quadratic form ($LGDP^2$); average years of schooling (SCH); log of exports share in GDP (LEXPGDP); continental dummy variables for Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM); service sector share in GDP (SERVGDP); closest air distance to a major port (AIRDIST); employment laws index (EMPLAW) and civil rights index (CIVILR). However, for simplicity we only report the coefficients on the measure of institutional quality and their corresponding statistics in the second stage and on the instrumental variables in the first stage.

A brief look at the second-stage results suggest that the relationship between our measure of institutional quality and the share of the small and medium enterprise sector is not robust to controlling for simultaneity bias. Further, the coefficient on institutional quality is of the wrong sign in Column 4, when agricultural transition year and absolute latitude are used as instruments. The Sargan tests of overidentification restrictions do not reject the null hypothesis that these instrument sets are valid.

Nevertheless, before drawing any conclusion, we need to determine the strength of the instrumental variables as weak instruments can bias our results and render conventional tests unreliable. To do so, we need to look at the F-statistics from the first stage of the instrumental variable estimation. Stock, Wright and Yogo (2002) suggest that, as a rule of thumb, values for the first-stage F-statistic below 10 would indicate a weak instrument problem. All of our first-stage F-statistics are well below this threshold suggesting that there exists a problem associated with weak instruments. In other words, the evidence fails to reject the null hypothesis that the instruments do not explain cross-country variation in the quality of

institutions. This is potentially due to the presence of explanatory variables that are highly correlated with the institutional variable in the estimated model, which makes precise estimation difficult. Three variables that are highly correlated with the institutional measure are log of income, its quadratic form and average years of schooling with correlations of 0.91, 0.92 and 0.74, respectively.

Table 5.10 presents results when the log of income per capita (LGDP), its quadratic form ($LGDP^2$) and average years of schooling (SCH) are dropped from the set of explanatory variables.[71]

Column 1 uses different legal origins as the instruments. Panel B indicates a strong first-stage relationship between legal origins and our measure of institutions. The second-stage result suggests that institutional quality has a positive effect on the share of the small and medium enterprise sector and the relationship is highly significant in all cases, except in Column 4 when agricultural transition year and absolute latitude are used as instruments. The Sargan test does not reject the null hypothesis that the instruments are valid. Furthermore, the F-statistic of 13.07 indicates that weak instruments is not an issue in this case.

In Column 2 and 3 we add (one at a time) agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) to legal origins and use them as instrument sets. First-stage results suggest that all of our instruments, except AGYEAR, have strong and positive relationships with institutional quality. On the other hand, AGYEAR has a significantly negative relationship though a weak one, based on the size of its coefficient. Second-stage results show that the addition of AGYEAR to the instrument sets in Column 2 slightly reduces the coefficient on our measure of institutions while the addition of ABSLATIT in Column 3 improves it. Further, the coefficient on institutional quality remains significantly different from zero with the additions of agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) to our set of instruments.

In Column 4, only AGYEAR and ABSLATIT are used as instruments. The first-stage relationship between AGYEAR and institutional quality remains weak though significant. The relationship between ABSLATIT and institutional quality, however, becomes insignif-

[71] The log of income per capita (LGDP), its quadratic form ($LGDP^2$) and average years of schooling (SCH) are dropped because, in addition to their strong correlations with the institutions variable, their relationship with the dependent variable is not statistically robust to changing set of conditioning variables, as shown in Table 5.6. Hence, their exclusion from the explanatory variable set may not pose any problem.

icant and weaker. Also, the second-stage relationship between our measure of institutions and the share of SME sector is positive but insignificant.

Finally, in Column 5, we include all instruments in our instrument sets. The use of this set of instruments does not have much effect on our previous findings. First-stage results indicate a strong relationship between each of the instruments with our measure of institutional quality and the signs on their respective coefficients remain the same. The second-stage result suggests that the relationship between institutional quality and the share of the SME sector is positive and significant. Furthermore, the p-value for the Sargan test is 0.726 suggesting that the evidence fails to reject the null hypothesis that our instruments are valid. Moreover, the first-stage F-statistic of 10.78 is slightly above our threshold of 10 implying that there is no issue with weak instrument.

Therefore, results from the instrumental variable estimation suggest a positive relationship between institutional quality and the share of small and medium enterprise sector though the relationship is not robust to changing set of instruments, i.e. it becomes insignificant when AGYEAR and ABSLATIT are used as instruments.

We experiment further with instrumental variable estimation with both institutional quality (INST) and the level of financial development (FINDEV) treated as endogenous. This is to take into account the possibility that our instruments have strong correlations with the financial development measure. Baum et al. (2003) point out that in the presence of multiple endogenous variables, commonly used statistics such as the F-test in the first-stage regression may not be sufficiently informative. To deal with this issue, the use of other statistics is required. One such statistic is the Shea Partial R^2 measure which was proposed by Shea (1997).[72] An advantage of this measure is that it takes into account the intercorrelations among the instruments. Moreover, instead of presenting the usual first-stage F-statistic, we will report a standard Partial R^2 of excluded instrument, which is the R^2 of the first-stage regression with the included instruments partialled-out.[73] As a rule of thumb, if an estimated equation produces a larger standard Partial R^2 than the Shea

[72] See Godfrey (1999) for a simplified method for calculating the Shea partial R^2.

[73] More specifically, this is the "squared partial correlation" between the excluded instruments and the endogenous regressor in question (Baum et al., 2003).

Partial R^2, we may conclude that the instruments lack sufficient relevance to explain all the endogenous regressors, and the model may be essentially unidentified (Baum et al., 2003).

The use of excluded instruments with little explanatory power would likely increase the bias in the estimated IV coefficients. Furthermore, if their explanatory power in the first-stage regression is nil, the model is in effect unidentified with respect to that endogenous variable. In such a case, instrumental variable estimation will become inconsistent and nothing is gained from instrumenting (Hahn and Hausman, 2002).

The results are presented in Table 5.11. Second-stage results suggest a positive but insignificant relationship between our institutional and financial variables on the one hand and the share of the SME sector on the other. The Sargan statistics fail to reject the null that the instruments are valid. However, first-stage regression statistics present a bleak picture. For all of the instrument sets, equation estimations yield large standard Partial R^2 and small Shea Partial R^2 values, which suggest that the models are unidentified and that our IV estimations are inconsistent. In order to deal with this issue, either additional relevant instruments are needed, or one of the endogenous regressors must be dropped from the model (Baum et al., 2003). The latter choice would result in the same estimation as in Table 5.10, which provides by far the most informative results.

Overall, instrumental variable estimation suggests that the relationship between our measure of the institutional quality and the employment share of the small and medium enterprise sector is not always robust to controlling for simultaneity bias. Therefore we can conclude that while countries with good institutions generally have a large SME sector, there is insufficient evidence to support convincingly the view that institutional quality exerts a causal effect on the size of the SME sector.

5.4 Methodology and Results: INST versus SEMSELE

We now determine potential disproportionate effects of institutions on small enterprise sectors versus large enterprise sectors. This relationship will show us whether the improvement in the quality of institutions works to benefit small enterprises more than it does for large enterprises, or vice versa. To do so, we use the UNIDO-based measure of the rela-

tive share of the small enterprise sector versus the large enterprise sector as the dependent variable. We follow the same procedure as we did above with the SME share as the dependent variable. Specifically, we first conduct OLS estimation of the relationship between our measure of institutions and the dependent variable; and then use instrumental variable estimation to investigate causality.

5.4.1 Ordinary Least Squares Estimation

In this case, the linear regression estimate takes the form:

$$SEMSELE_i = \alpha + \beta\, INST_i + \gamma\, X_i + u$$

where $SEMSELE$ is the relative size of the small enterprise sector versus the large enterprise sector, $INST$ is the institutional quality indicator measured, X is a set of conditioning variables, i is the country index and u is the error term.

Table 5.12 displays the results of the Ordinary Least Squares regression estimation. We begin by using institutional quality as the only explanatory variable in Column 1. The result in this case shows a slightly negative simple correlation between institutional quality and the dependent variable. Figure 5.4 shows this simple relationship.

In Columns 2-6, we add the log of income per capita and its quadratic form, average years of schooling, log of share of exports in GDP, dummies for different continents, service sector share in GDP, air distance to a major port, civil rights index, employment law index and the level of financial development to our conditioning variable set. The results suggest that, conditioning on other variables, our measure of institutional quality has a positive correlation with the dependent variable. The relationships are, however, not significant at the conventional levels. Figure 5.5 shows the partial relationship between the measure of institutional quality and the dependent variable for our sample of 57 countries. In addition, R^2 values suggest that our independent variables do not have much explanatory power for the cross-country variation in the relative share of the small enterprise sector versus its larger counterpart.[74]

[74] The best R^2 value is only about 0.227 suggesting that our independent variables explain less than a quarter

Table 5.13 shows the relationship between each of the six components of our institutional measure and the dependent variable. Coefficient estimates and conventional test statistics show that four of the six variables have positive but insignificant relationships with the dependent variable. Two variables – regulatory quality (REGQTY) and government effectiveness (GOVEFF) – have negative but insignificant coefficient estimates when log of income and its quadratic form are included in the list of conditioning variables. These relationships turn positive but remains insignificant when log of income and its quadratic form are not used as conditioning variables. Figure 5.6 show the corresponding partial relationships.

However, as explained earlier, the OLS results should be viewed with caution because they might be subject to problems such as simultaneity bias and measurement error.

5.4.2 Instrumental Variable Estimation

We now examine this relationship further through the use of instrumental variables. We use the same five different instrument sets as in the previous analysis. They include the different legal origins (French, UK, German and Scandinavian legal origins), agriculture transition year (AGYEAR) and absolute latitude (ABSLATIT).

Sargan test statistics in Panel A of Table 5.14 again fail to reject the null that the instruments are valid. Second-stage results indicate a positive but insignificant relationship between institutional quality and the dependent variable in the first two columns when only legal origins and agricultural years are used as instruments. In Columns 3-5, this relationship becomes negative but still remains insignificant. Furthermore, first-stage F-statistics in Panel B suggest a potential weak instrument problem. Table 5.15 excludes the log of income per capita, its quadratic form and average years of schooling from the conditioning variable set. Only the set of instruments shown in Column 4, which comprise the agricultural transition year and absolute latitude, does not suffer from a potential weak instrument issue with an F-statistic of 22.56. The Sargan test statistics is at 0.317, which is well above rejection level. Moreover, the coefficient on our variable of interest has a negative sign though the relationship is insignificant.

---

of the cross-country variation in the relative share of the small enterprise sector versus its larger counterpart.

Finally, in Table 5.16 we treat both the measure of institutions and the level of financial development as endogenous. Second-stage results from the table suggest a positive but insignificant association between the dependent variable and our measure of institutional quality in four of five cases. It turns negative, however, when all instruments are used at the same time. Similarly, relationship between financial system development and the dependent variable is unclear with, indications of a positive association in two cases and a negative association in others. Further, all of these relationships are insignificant. Moreover, first-stage regression estimations yield large standard Partial R^2 compared to Shea Partial R^2 values suggesting that the models are unidentified and that our instrumental variable estimation may be inconsistent.

Therefore, taking into account various relevant statistics, the regression estimates in Table 5.15 are likely to provide the most informative results. In this case, instrumental variable estimation does not show any clear disproportionate effect of institutional quality on the relative size of the small enterprise sector versus the large enterprise sector.

5.5 Conclusion

In this paper we seek to explore the potential effects of institutions on the firm size distribution. This is done in two different ways. First, a causal relationship between institutional quality and the small and medium enterprise employment share in the manufacturing sector is examined. Second, potential disproportionate effects of institutions on firms of different sizes are investigated by examining the relationship between the measure of institutional quality and the relative share of the small enterprise sector versus the large enterprise sector.

Instrumental variable methods are employed in our analysis to determine the causality of the relationship. The results indicate that although our measure of institutions has a positive relationship with the share of the SME sector, this relationship is not robust to controlling for simultaneity bias. This suggests that while countries with good institutions generally have a large SME sector, there is insufficient evidence to convincingly support the view that institutional quality exerts a causal effect on the size of the SME sector.

In addition, the relationship between institutional quality and the relative size of the small enterprise sector versus the large enterprise sector is even less clear. The results from both Ordinary Least Squares and the instrumental variable estimation do not show any strong correlation between them.

The results, however, do not allow a firm conclusion that institutions have no effect on the firm size distribution. Given the potentially important role of this relationship in the economic growth and development process, further investigation into this relationship may be useful, perhaps through the use of different measures of both institutions and the SME sector, or a more informative set of instruments.

Table 5.1a: List of Countries (when SMEOFF is the dependent variable)

ALB Albania	KAZ Kazakhstan
ARG Argentina	KEN Kenya
AUS Australia	KOR Korea
AUT Austria	KGZ Kyrgyz Rep
AZE Azerbaijan	LVA Latvia
BLR Belarus	LUX Luxembourg
BEL Belgium	MEX Mexico
BRA Brazil	NLD Netherlands
BRN Brunei	NZL New Zealand
BGR Bulgaria	NIC Nicaragua
BDI Burundi	NGA Nigeria
CMR Cameroon	NOR Norway
CAN Canada	PAN Panama
CHL Chile	PER Peru
COL Colombia	PHL Philippines
CRI Costa Rica	POL Poland
CIV Cote D'Ivoire	PRT Portugal
HRV Croatia	ROM Romania
CZE Czech Republic	RUS Russian
DNK Denmark	SGP Singapore
ECU Ecuador	SVK Slovak Rep
SLV El Salvador	SVN Slovenia
EST Estonia	ZAF South Africa
FIN Finland	ESP Spain
FRA France	SWE Sweden
GEO Georgia	CHE Switzerland
DEU Germany	TWN Taiwan
GHA Ghana	TJK Tajikistan
GRC Greece	TZA Tanzania
GTM Guatemala	THA Thailand
HND Honduras	TUR Turkey
HKG Hong Kong	UKR Ukraine
HUN Hungary	GBR United Kingdom
ISL Iceland	USA USA
IDN Indonesia	VNM Vietnam
IRL Ireland	YUG Yugoslavia
ITA Italy	ZMB Zambia
JPN Japan	ZWE Zimbabwe

Notes: This table shows names and country codes for the 76 sample countries used for our analysis, with the share of the small and medium enterprise sector (SMEOFF) is used as the dependent variable. They are listed in alphabetical order.

Table 5.1b: List of Countries (when SEMSELE is the dependent variable)

ARG Argentina	MAR Morocco
AUS Australia	NLD Netherlands
AUT Austria	NZL New Zealand
BGD Bangladesh	NGA Nigeria
BOL Bolivia	NOR Norway
BGR Bulgaria	PAK Pakistan
CMR Cameroon	PAN Panama
CAN Canada	PER Peru
CHL Chile	PHL Philippines
COL Colombia	POL Poland
CRI Costa Rica	PRT Portugal
CYP Cyprus	SGP Singapore
DNK Denmark	ESP Spain
ECU Ecuador	LKA Sri Lanka
EGY Egypt	SWE Sweden
SLV El Salvador	THA Thailand
FIN Finland	TUR Turkey
FRA France	GBR United Kingdom
GHA Ghana	VEN Venezuela
GRC Greece	ZWE Zimbabwe
GTM Guatemala	
HUN Hungary	
ISL Iceland	
IND India	
IDN Indonesia	
IRN Iran	
IRL Ireland	
ISR Israel	
ITA Italy	
JPN Japan	
JOR Jordan	
KEN Kenya	
KOR Korea	
KWT Kuwait	
MYS Malaysia	
MLT Malta	
MUS Mauritius	

Notes: This table shows names and country codes for the 57 sample countries used for our analysis, with the relative share of the small enterprise sector versus the large enterprise sector (SEMSELE) is used as the dependent variable. They are listed in alphabetical order.

Table 5.2: Correlation between Institutional Quality Measure and Various Indexes of Governance

	INST	VACC	PSTAB	REGQTY	RLAW	GOVEFF	CCORR
INST	1.00						
VACC	0.92	1.00					
PSTAB	0.88	0.80	1.00				
REGQTY	0.91	0.78	0.77	1.00			
RLAW	0.98	0.90	0.81	0.87	1.00		
GOVEFF	0.97	0.85	0.80	0.89	0.96	1.00	
CCORR	0.96	0.88	0.80	0.83	0.96	0.95	1.00

Notes: This table presents the correlation between the aggregate measure of institutional quality on the one hand and each of the six indexes of governance on the other. “VACC” stands for voice and accountability, “PSTAB” for political stability, “REGQTY” for regulatory quality, “RLAW” for rule of law, “GOVEFF” for government effectiveness, and, finally, “CCORR” for control of corruption.

Table 5.3: Variable Description and Sources

VARIABLES	DESCRIPTION	SOURCES
SMEOFF	Share of small and medium enterprises in manufacturing sector, when official country definition of SMEs is used.	Ayyagari, Beck and Demirguc-Kunt (2003) new database on SMEs
SEMSELE	The relative share of the small enterprise sector versus the large enterprise sector.	The data used to calculate SEMSELE are from UNIDO Industrial Statistics Database (2005)
INST	A measure of institutional quality. It is the average of six different dimensions of governance: voice and accountability, political stability and violent, government effectiveness, regulatory quality, rule of law, and control of corruption.	Kaufmann et al. (1999)
LGDP	Log of real gross domestic per capita, measured in constant price, averaged over 1990-1999.	Summers, Heston and Aten's PWT version 6.1
$LGDP^2$	Square of log of real gross domestic per capita, measured in constant price, averaged over 1990-1999.	Ibid.
SCH	Average years of schooling for the population age 15 and over.	Barro and Lee (2001) Educational Attainment Dataset, updated version
LEXPGDP	Log of exports as a percentage of GDP.	World Bank World Development Indicators (2006)
ESEASIA	Dummy for East Asian countries.	Harvard University's Center for International Development (CID) geography dataset
EU	Dummy for European countries.	Ibid.
SAFRI	Dummy for Sub-Saharan African countries.	Ibid.
LATAM	Dummy for Latin American countries.	Ibid.
SERVGDP	Share of the service sector in GDP.	World Bank World Development Indicators (2006).
AIRDIST	The closest distance in kilometers to a major port.	Harvard University's CID geography dataset
CIVILR	Civil rights index which assesses the degree of protection of vulnerable groups against employment discrimination.	La Porta et al. (2005) Regulation of Labor data
EMPLAW	Employment laws index measure of the protection of labor and employment laws.	Ibid.

VARIABLES	DESCRIPTION	SOURCES
BANK	The ratio of bank credits divided by bank credits plus central bank domestic assets.	Beck, Levine and Loayza (2000) financial dataset
LLY	Liquid liabilities – which equals to liquid liabilities of financial system (currency plus demand and interest-bearing liabilities of banks and nonblank financial intermediaries) divided by GDP.	Ibid.
PRIVATE	The ratio of private credits by deposit money banks and other financial institutions to GDP.	Ibid.
FINDEV	The measure of the level of financial development. It is the first principal component of BANK, LLY and PRIVATE.	Ibid.
French legal origin	French legal origin. It indicates countries whose legal system is based on the French legal system.	La Porta et al. (1997)
UK legal origin	UK legal origin. It indicates countries whose legal system is influenced by the British legal system.	Ibid.
German legal origin	German legal origin. It indicates countries whose legal system is influenced by the German legal system.	Ibid.
Scandinavian legal origin	Scandinavian legal origin. It indicates countries whose legal system is influenced by the Scandinavian legal system.	Ibid.
AGYEAR	Agricultural transition year, indicating the estimate of the year when a country made the transition to an agrarian state.	Putterman, L. (2006) Data can be obtained at the following website: *http://www.econ.brown.edu/fac/Louis_Putterman/*
ABSLATIT	Absolute latitude or absolute distance from the equator.	World Bank Global Development Network Growth Database

Table 5.4: Descriptive Statistics

Variable	Observation	Mean	Standard Deviation	Minimum	Maximum	25th Percentile	Median	75th Percentile
SEMSELE	41	53.76	10.85	36.67	82.51	45.75	50.73	60.88
SMEOFF	41	59.39	17.48	15.20	86.70	51.61	61.05	72.10
SMEOFF	76	51.44	22.70	4.59	86.70	33.60	58.54	69.32
INST	76	0.40	0.98	-1.75	1.93	-0.36	0.24	1.34
VACC	76	0.34	0.99	-1.49	1.76	-0.40	0.35	1.34
PSTAB	76	0.29	0.92	-2.67	1.59	-0.36	0.41	1.07
GOVEFF	76	0.49	1.13	-1.47	2.51	-0.46	0.24	1.67
REGQTY	76	0.52	0.96	-1.88	2.58	-0.15	0.53	1.39
RLAW	76	0.43	1.09	-1.41	2.17	-0.49	0.27	1.65
CCORR	76	0.34	1.13	-1.64	2.24	-0.70	0.06	1.33
LGDP	76	8.86	1.05	6.16	10.47	8.26	8.88	9.89
$LGDP^2$	76	79.65	18.05	37.89	109.56	68.27	78.95	97.86
SCH	76	7.51	2.41	1.38	11.89	5.62	7.79	9.36
LEXPGDP	76	3.46	0.59	2.16	5.21	3.17	3.51	3.73
ESEASIA	76	0.13	0.34	0	1	0	0	0
EU	76	0.24	0.43	0	1	0	0	0
SAFRI	76	0.13	0.34	0	1	0	0	0
LATAM	76	0.17	0.38	0	1	0	0	0
SERVGDP	76	54.56	12.49	22.54	82.71	47.50	55.56	64.61
AIRDIST	76	3268	2555	140	9280	1150	2703	5230

Variable	Observation	Mean	Standard Deviation	Minimum	Maximum	25th Percentile	Median	75th Percentile
CIVILR	76	0.69	0.12	0.23	0.93	0.59	0.72	0.78
EMPLAW	76	0.49	0.18	0.15	0.83	0.36	0.48	0.66
FINDEV	76	0.00	0.95	-1.40	2.91	-0.67	-0.31	0.65
UK Legal Origin	76	0.22	0.42	0	1	0	0	0
French Legal Origin	76	0.36	0.48	0	1	0	0	1
German Legal Origin	76	0.08	0.27	0	1	0	0	0
Scand. Legal Origin	76	0.07	0.25	0	1	0	0	0
ABSLATIT	76	32.98	18.79	0.51	63.89	13.85	37.93	48.32
AGYEAR	74	4882	1992	400	10000	3500	5000	6500

Table 5.5: Correlation between variables

	SMEOFF	SEMSELE	INST	LGDP	LGDP2	SCH	LEXPGDP	ESEASIA	EU	SAFRI	LATAM	SERVGDP	AIRDIST	CIVILR	EMPLAW	FINDEV
SMEOFF	1.00															
SEMSELE	0.02	1.00														
INST	0.41	-0.10	1.00													
LGDP	0.49	-0.24	0.91	1.00												
$LGDP^2$	0.47	-0.23	0.92	0.99	1.00											
SCH	0.34	-0.23	0.74	0.74	0.75	1.00										
LEXPGDP	-0.13	-0.12	0.29	0.13	0.15	0.11	1.00									
ESEASIA	0.29	-0.27	-0.03	0.03	0.03	-0.01	0.22	1.00								
EU	0.27	-0.01	0.62	0.60	0.62	0.32	0.14	-0.29	1.00							
SAFRI	-0.69	0.22	-0.54	-0.70	-0.67	-0.53	0.00	-0.15	-0.26	1.00						
LATAM	0.07	0.27	-0.34	-0.27	-0.30	-0.28	-0.33	-0.21	-0.38	-0.19	1.00					
SERVGDP	0.39	-0.14	0.73	0.79	0.78	0.58	-0.05	-0.21	0.49	-0.62	0.04	1.00				
AIRDIST	-0.25	0.34	-0.38	-0.46	-0.46	-0.33	-0.15	0.00	-0.59	0.38	0.34	-0.30	1.00			
CIVILR	0.02	0.08	-0.28	-0.25	-0.27	-0.17	-0.47	-0.46	-0.12	0.07	0.27	-0.02	-0.07	1.00		
EMPLAW	0.39	-0.06	0.25	0.29	0.28	0.13	0.09	-0.12	0.59	-0.34	-0.12	0.32	-0.42	0.01	1.00	
FINDEV	0.36	-0.32	0.69	0.71	0.72	0.54	0.22	0.43	0.36	-0.37	-0.40	0.50	-0.37	-0.28	0.07	1.00

Table 5.6: Ordinary Least Squares (OLS) Regression Estimation with SMEOFF as the Dependent Variable

Independent Variable	(1)	(2)	(3)	(4)	(5)	(6)
CONSTANT	46.104 (19.778)	-155.203 (-1.100)	-155.115 (-1.150)	-188.024 (-1.290)	-251.475 (-1.680)	-241.783 (-1.550)
INST	13.265** (5.981)	16.180** (2.880)	15.682** (2.2890)	15.170*** (3.040)	12.913** (2.400)	13.127** (2.390)
LGDP		49.685 (1.410)	53.784 (1.600)	64.752* (1.850)	75.906** (2.010)	73.080* (1.840)
$LGDP^2$		-3.016 (-1.380)	-2.892 (-1.380)	-3.951* (-1.850)	-4.722** (-2.020)	-4.536* (-1.830)
SCH			-2.850** (-2.340)	-0.383 (-0.280)	-0.179 (-0.130)	-0.174 (-0.130)
LEXPGDP			-7.144* (-1.930)	-9.876*** (-2.820)	-10.159*** (-2.790)	-10.185*** (-2.770)
ESEASIA				32.371*** (4.670)	34.378*** (4.460)	35.643*** (3.810)
EU				17.685 (2.570)**	25.507*** (3.230)	25.782*** (3.210)
SAFRI				2.951 (0.320)	-5.336 (-0.480)	-5.046 (-0.450)
LATAM				13.268** (2.010)	5.989 (0.720)	6.108 (0.720)
SERVGDP					0.207 (0.670)	0.217 (0.700)
AIRDIST					0.002 (1.420)	0.002 (1.380)
CIVILR					26.598 (1.360)	26.692 (1.360)
EMPLAW					-16.582 (-1.260)	-17.211 (-1.280)
FINDEV						-1.105 (-0.240)
R^2	0.326	0.344	0.421	0.579	0.618	0.618
N	76	76	76	76	76	75

Notes: The dependent variable for the regression is the share of small and medium enterprise in the manufacturing sector (SMEOFF). This table shows the OLS regression of the share of small and medium enterprise on our measure of institutional quality (INST) and a number of conditioning variables. The numbers in parentheses are the t-statistics. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Table 5.7: OLS Regression Estimations of SMEOFF on Different Components of Institutional Quality

Independent Variable	**Panel A: Including log of income and its quadratic form in the regression**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
VACC	9.549** (2.210)					
PSTAB		3.159 (0.870)				
REGQTY			13.256*** (3.160)			
RLAW				13.152** (2.420)		
GOVEFF					11.559** (2.220)	
CCORR						4.002 (0.940)
R^2	0.614	0.587	0.642	0.619	0.614	0.588
N	75	75	75	75	75	75
	Panel B: Excluding log of income and its quadratic form from the regression					
VACC	9.318*** (2.280)					
PSTAB		3.104 (0.900)				
REGQTY			11.453*** (2.980)			
RLAW				8.973* (1.920)		
GOVEFF					8.888** (2.040)	
CCORR						3.498 (0.940)
R^2	0.600	0.572	0.621	0.591	0.594	0.573
N	75	75	75	75	75	75

Note: The dependent variable is the share of small and medium enterprises in the manufacturing sector (SMEOFF). This table shows the relationship between each component of our measure of institutional quality and the dependent variable. All regressions in Panel A include a set of conditioning variables comprising of log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average years of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), share of service sector in GDP (SERVGDP), air distant (AIRDIST), civil rights index (CIVILR), employment laws index (EMPLAW), and measure of financial development (FINDEV) . However, the estimates on these conditioning variables are not reported in the table in order to save space. In Panel B, we drop log of income per capita and its quadratic form from our conditioning set of variables. The numbers in parentheses are the t-statistics. *, **, *** indicate significance levels of 10, 5 and 1 percent, respectively.

Table 5.8: Extreme Bound Analysis (EBA) Results

Variable of Interest	B	t-stat	Observation	R2	Other Variables	Robust/Fragile
Panel A: EBA when LGDP and SCH are included in the base regression						
INST	High: 15.482	3.099	76	0.513	LEXPGDP,ESEASIA, SAFRI,FINDEV	Robust
	Base: 10.731	2.360	76	0.376		
	Low: 9.423	2.013	75	0.516	LEXPGDP,ESEASIA, EU,AIRDIST	
Panel B: EBA when LGDP and SCH are excluded from the base regression						
INST	High: 17.647	4.512	76	0.487	SCH,LEXPGDP,ESEASIA,FINDEV	Robust
	Base: 13.265	5.980	76	0.326		
	Low: 7.178	2.073	76	0.437	ESEASIA,SAFRI,SERVGDP,EMPLAW	

Notes: The dependent variable is the share of the small and medium enterprises in the manufacturing sector (SMEOFF). This table shows test results using extreme bound analysis on the relationship between our measure of institutional quality and SMEOFF. The Base B in Panel A is the estimated coefficient on INST from the base regression in which the variable of interest (INST), log of income per capita (LGDP) and average years of schooling (SCH) are used as explanatory variables. The Base in Panel B is the estimated coefficient on INST from the regression in which only the variable of interest, INST, is included. The High B is the estimated coefficient on INST from the regression with the high extreme bound, while the Low B is the estimated coefficient on INST from the regression with the extreme low bound. The "Other Variables" are the combination of conditioning variables which produce the extreme bounds. The "Robust/Fragile" indicates whether the variable of interest, INST, is robust or fragile.

Table 5.9: IV Estimation of the Relationship between Institutional Quality and the Share of the SME Sector

	Panel A: Second-Stage Regression (Dependent Variable: Share of Small and Medium Enterprise)				
	(1)	(2)	(3)	(4)	(5)
INST	4.626 (0.200)	1.268 (0.090)	4.699 (0.210)	-4.666 (-0.250)	1.374 (0.090)
Sargan test (p-value)	0.936	0.613	0.979	0.699	0.613
Observations	75	73	75	73	73
	Panel B: First-Stage Regression (Dependent Variable: Measure of Institutional Quality)				
UK legal origin	0.327 (1.130)	0.087 (0.290)	0.332 (1.060)		0.093 (0.280)
French legal origin	0.175 (0.700)	0.041 (0.170)	0.179 (0.660)		0.046 (0.170)
German legal origin	0.409 (1.230)	0.279 (0.850)	0.412 (1.210)		0.282 (0.830)
Scandinavian legal origin	0.511 (1.500)	0.418 (1.170)	0.511 (1.480)		0.419 (1.160)
AGYEAR		-0.0001* (-1.830)		-0.0001** (-2.200)	-0.0001* (-1.800)
ABSLATIT			0.0003 (0.04)	0.003 (0.440)	0.0003 (0.050)
F-test	0.72	1.48	0.57	2.48	1.21
Observations	75	73	75	73	73
R^2	0.887	0.895	0.887	0.891	0.895

Notes: this table presents instrumental variable (IV) regression estimates for the cross-section of countries. Panel A reports the second-stage regression estimates from IV regressions with the share of small and medium enterprise sector (SMEOFF) as the dependent variable. "Sargan test" is the test of over-identification restrictions and is used to test the validity of the instruments. The numbers in parentheses are the z-statistics. Panel B presents the first-stage regression estimates with the measure of institutional quality (INST) as the dependent variable. "F-test" is the test for excluded instruments. The numbers in parentheses are the t-statistics. Conditioning variables used in both first-stage and second-stage regressions include log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average year of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), air distant (AIRDIST), employment laws index (EMPLAW) and civil rights index (CIVILR). We also include the number of observation and R^2 value for regressions in both first and second stages. Column 1 uses different legal origins -- UK legal origin, French legal origin, German legal origin and Scandinavian legal origin -- as instruments; Column 2 uses legal origins and agricultural transition year (AGYEAR); Column 3 uses legal origins and absolute latitude (ABSLATIT); Column 4 uses agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) and Column 5 uses legal origins, agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) as instruments.
*, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Table 5.10: IV Estimation of the Relationship between Institutional Quality and the Share of the SME Sector

	Panel A: Second-Stage Regression (Dependent Variable: Share of Small and Medium Enterprise)				
	(1)	(2)	(3)	(4)	(5)
INST	10.838*** (2.870)	10.719*** (2.900)	11.387*** (3.190)	7.006 (1.410)	11.615*** (3.290)
Sargan test (p-value)	0.753	0.611	0.844	0.557	0.726
Observations	75	73	75	73	73
	Panel B: First-Stage Regression (Dependent Variable: Measure of Institutional Quality)				
UK legal origin	1.813*** (6.710)	1.438*** (4.610)	1.938*** (7.360)		1.627*** (5.200)
French legal origin	1.177*** (3.780)	0.975*** (3.100)	1.349*** (4.410)		1.167*** (3.700)
German legal origin	1.958*** (5.110)	1.656*** (4.180)	1.859*** (5.040)		1.630*** (4.260)
Scandinavian legal origin	1.862*** (4.410)	1.515*** (3.190)	1.657*** (4.020)		1.444*** (3.140)
AGYEAR		-0.0001* (-1.970)		-0.0002*** (-4.550)	-0.0001* (-1.530)
ABSLATIT			0.021** (2.580)	0.010 (1.210)	0.019** (2.290)
F-test	13.07	11.10	12.72	12.25	10.78
Observations	75	73	75	73	73
R^2	0.743	0.758	0.768	0.664	0.778

Notes: this table presents instrumental variable (IV) regression estimates for the cross-section of countries. This table differs from table 9 in that it excludes the log of income per capita (LGDP), its quadratic form ($LGDP^2$) and the average years of schooling (SCH) from the explanatory variable list. Panel A reports the second-stage regression estimates from IV regressions with the share of small and medium enterprise sector (SMEOFF) as the dependent variable. "Sargan test" is the test of over-identification restrictions and is used to test the validity of the instruments. The numbers in parentheses are the z-statistics. Panel B presents the first-stage regression estimates with the measure of institutional quality (INST) as the dependent variable. "F-test" is the test for excluded instruments. The numbers in parentheses are the t-statistics. Conditioning variables used in both first-stage and second-stage regressions include log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), air distant (AIRDIST), employment laws index (EMPLAW) and civil rights index (CIVILR). We also include the number of observation and R^2 value for regressions in both first and second stages. Column 1 uses different legal origins -- UK legal origin, French legal origin, German legal origin and Scandinavian legal origin -- as instruments; Column 2 uses legal origins and agricultural transition year (AGYEAR); Column 3 uses legal origins and absolute latitude (ABSLATIT); Column 4 uses agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) and Column 5 uses legal origins, agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) as instruments. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Table 5.11: IV Estimation of the Relationship between Institutional Quality and the Share of the SME Sector

Panel A: Second-Stage Regression
(Dependent Variable: Share of Small and Medium Enterprises in the Manufacturing Sector, SMEOFF)

	(1)	(2)	(3)	(4)	(5)
INST	4.283 (0.460)	4.906 (0.520)	6.937 (0.840)	21.072 (0.760)	8.353 (0.990)
FINDEV	7.514 (0.780)	6.783 (0.680)	5.384 (0.600)	-21.121 (-0.520)	3.961 (0.431)
Sargan test (p-value)	0.777	0.695	0.806	-	0.638
Observations	75	73	75	73	73

Panel B: First-Stage Regression
(Dependent Variable: Measure of Institutional Quality (INST) and the Level of Financial Development (FINDEV))

	INST	FINDEV	INST	FINDEV	INST	FINDEV	INST	FINDEV	INST	FINDEV
UK legal origin	1.812*** (6.710)	1.502*** (5.900)	1.438*** (4.610)	1.330*** (4.380)	1.937*** (7.360)	1.556*** (6.020)			1.627*** (5.200)	1.420*** (4.490)
French legal origin	1.177*** (3.780)	0.807*** (2.750)	0.975*** (3.100)	0.722*** (2.350)	1.349*** (4.410)	0.881*** (2.940)			1.167*** (3.700)	0.807** (2.530)
German legal origin	1.957*** (5.110)	1.865*** (5.170)	1.656*** (4.180)	1.742*** (4.500)	1.859*** (5.040)	1.823*** (5.030)			1.630*** (4.260)	1.730*** (4.470)
Scandinavian legal origin	1.862*** (4.410)	0.770* (1.930)	1.514*** (3.190)	0.681 (1.470)	1.657*** (4.020)	0.682* (1.680)			1.443*** (3.140)	0.649 (1.400)
AGYEAR			-0.0001** (-1.970)	-0.0001 (-0.950)			-0.0002*** (-4.550)	-0.0001*** (-3.140)	-0.0001 (-1.530)	-0.00004 (-0.730)
ABSLATIT					0.021** (2.580)	0.009 (1.130)	0.010 (1.210)	0.001 (0.140)	0.019** (2.290)	0.008 (1.000)
Partial R^2	0.453	0.463	0.481	0.461	0.506	0.475	0.280	0.141	0.523	0.470
Shea Partial R^2	0.161	0.165	0.163	0.156	0.198	0.185	0.023	0.012	0.189	0.170
Observation	75	75	73	73	75	75	73	73	73	73
Joint Sig Test (p-value)	0.149		0.207		0.141		0.439		0.154	

Notes: this table presents instrumental variable (IV) regression estimates for the cross-section of countries. Panel A reports the second-stage regression estimates from IV regressions with the share of small and medium enterprise sector (SMEOFF) as the dependent variable. "Sargan test" is the test of over-identification restrictions and is used to test the validity of the instruments. The numbers in parentheses are the standard errors. Panel B presents the first-stage regression estimates with the measure of institutional quality (INST) as the dependent variable. "Partial R^2" is the R^2 of the first-stage regression with included instruments partialled-out. "Shea Partial R^2" is a partial R^2 that takes into account the intercorrelations among the instruments. "Joint Sig Test" is the test of joint significance of endogenous regressors. The numbers in parentheses are the t-statistics. Conditioning variables used in both first-stage and second-stage regressions include log of share of exports in GDP (LEXPGDP); continental dummy variables for Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM); share of service sector in GDP (SERVGDP); closest distance to a major port (AIRDIST); civil rights index (CIVILR) and employment law index (EMPLAW). We also include the number of observation and R^2 value for regressions in both first and second stage. Column 1 uses different legal origins -- UK legal origin, French legal origin, German legal origin and Scandinavian legal origin -- as instruments; Column 2 uses legal origins and agricultural transition year (AGYEAR); Column 3 uses legal origins and absolute latitude (ABSLATIT); Column 4 uses agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) and Column 5 uses legal origins, agricultural transition year (AGYEAR), and absolute latitude (ABSLATIT) as instruments. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Table 5.12: Ordinary Least Squares (OLS) Regressions for SEMSELE

Independent Variable	(1)	(2)	(3)	(4)	(5)	(6)
CONSTANT	54.247 (30.040)	136.318 (0.880)	110.230 (0.690)	161.092 (0.860)	195.969 (0.980)	203.173 (0.990)
INST	-0.157 (-0.090)	1.263 (0.300)	1.776 (0.380)	2.819 (0.610)	2.024 (0.390)	2.011 (0.390)
LGDP		-17.305 (-0.470)	-12.613 (-0.330)	-27.676 (-0.630)	-34.997 (-0.730)	-37.453 (-0.750)
$LGDP^2$		0.890 (0.410)	0.651 (0.300)	1.643 (0.640)	2.063 (0.740)	2.216 (0.770)
SCH			-0.698 (-0.590)	-0.811 (-0.700)	-0.662 (-0.540)	-0.694 (-0.560)
LEXPGDP			2.424 (0.800)	3.598 (1.150)	2.475 (0.720)	2.673 (0.750)
ESEASIA				-8.605 (-1.610)	-9.644 (-1.570)	-8.756 (-1.190)
EU				-4.579 (-0.900)	-5.938 (-0.890)	-5.919 (-0.880)
SAFRI				5.753 (0.820)	7.105 (0.910)	6.883 (0.870)
LATAM				8.656* (1.830)	8.946* (1.770)	8.580 (1.600)
SERVGDP					0.105 (0.370)	0.139 (0.430)
AIRDIST					0.001 (0.140)	0.001 (0.130)
CIVILR					-15.046 (-1.080)	-14.648 (-1.030)
EMPLAW					7.303 (0.580)	6.816 (0.530)
FINDEV						-0.804 (-0.230)
R^2	0.0001	0.010	0.031	0.194	0.226	0.227
N	57	57	57	57	57	57

Notes: The dependent variable for the regression is the relative size of the small enterprise sector versus the large enterprise sector (SEMSELE). This table shows the OLS regression of the relative size of the small and medium enterprise sector versus the large enterprise sector on our measure of institutional quality (INST) and a number of conditioning variables. The numbers in parentheses are the t-statistics. *, ** indicate significance level of 10 percent and 5 percent, respectively.

Table 5.13: OLS regression the relative share of the small enterprise sector versus the large enterprise sector (SEMSELE) on different measures of institutional quality

Independent Variable	**Panel A: Including log of income and its quadratic form in the regression**					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
VACC	0.683 (0.150)					
PSTAB		2.792 (0.800)				
REGQTY			-0.221 (-0.060)			
RLAW				0.714 (0.170)		
GOVEFF					-0.549 (-0.12)	
CCORR						2.523 (0.690)
R^2	0.224	0.235	0.224	0.225	0.224	0.233
N	57	57	57	57	57	57
	Panel B: Excluding log of income and its quadratic form from the regression					
VACC	1.931 (0.450)					
PSTAB		3.223 (0.980)				
REGQTY			1.437 (0.040)			
RLAW				2.355 (0.670)		
GOVEFF					1.573 (0.400)	
CCORR						3.394 (1.130)
R^2	0.207	0.219	0.206	0.210	0.206	0.225
N	57	57	57	57	57	57

Note: The dependent variable for the regression is the relative size of the small enterprise sector versus the large enterprise sector (SEMSELE). This table shows the relationship between each component of our measure of institutional quality and the dependent variable. All regressions in Panel A include a set of conditioning variables comprising of log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average years of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), share of service sector in GDP (SERVGDP), air distant (AIRDIST), civil rights index (CIVILR), employment laws index (EMPLAW), and measure of financial development (FINDEV) . However, the estimates on these conditioning variables are not reported in the table in order to save space. In Panel B, we drop log of income per capita and its quadratic form from our conditioning set of variables. The numbers in parentheses are the t-statistics. *, **, *** indicate significance levels of 10, 5 and 1 percent, respectively.

Table 5.14: IV Estimation of the Relationship between Institutional Quality and the Relative Size of the Small Enterprise Sector

	Panel A: Second-Stage Regression (Dependent Variable: Share of Small and Medium Enterprise)				
	(1)	(2)	(3)	(4)	(5)
INST	0.286 (0.02)	3.903 (0.430)	-5.179 (-0.500)	-1.422 (-0.170)	-2.447 (-0.340)
Sargan test (p-value)	0.315	0.827	0.454	0.218	0.753
Observations	57	55	57	55	55
	Panel B: First-Stage Regression (Dependent Variable: Measure of Institutional Quality)				
UK legal origin	0.451* (1.900)	0.316 (1.330)	0.570** (2.450)		0.438* (1.950)
French legal origin	0.234 (0.990)	0.128 (0.550)	0.301 (1.320)		0.209 (0.950)
German legal origin	0.214 (0.630)	0.144 (0.430)	0.190 (0.590)		0.128 (0.410)
Scandinavian legal origin	0.344 (1.080)	0.257 (0.760)	0.257 (0.840)		0.199 (0.640)
AGYEAR		-0.00006** (-2.240)		-0.0001*** (-3.190)	-0.00001* (-2.690)
ABSLATIT			0.015** (2.200)	0.014** (2.220)	0.018** (2.640)
F-test	1.04	1.89	1.88	6.57	3.00
Observations	57	55	57	55	55
R^2	0.894	0.906	0.906	0.911	0.921

Notes: this table presents instrumental variable (IV) regression estimates for the cross-section of countries. Panel A reports the second-stage regression estimates from IV regressions with the relative size of the small enterprise sector versus the large enterprise sector (SEMSELE) as the dependent variable. "Sargan test" is the test of over-identification restrictions and is used to test the validity of the instruments. The numbers in parentheses are the z-statistics. Panel B presents the first-stage regression estimates with the measure of institutional quality (INST) as the dependent variable. "F-test" is the test for excluded instruments. The numbers in parentheses are the t-statistics. Conditioning variables used in both first-stage and second-stage regressions include log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average year of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), air distant (AIRDIST), employment laws index (EMPLAW) and civil rights index (CIVILR). We also include the number of observation and R^2 value for regressions in both first and second stages. Column 1 uses different legal origins -- UK legal origin, French legal origin, German legal origin and Scandinavian legal origin -- as instruments; Column 2 uses legal origins and agricultural transition year (AGYEAR); Column 3 uses legal origins and absolute latitude (ABSLATIT); Column 4 uses agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) and Column 5 uses legal origins, agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) as instruments. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Table 5.15: IV Estimation of the Relationship between Institutional Quality and the Relative Size of the Small Enterprise Sector

	Panel A: Second-Stage Regression (Dependent Variable: Share of Small and Medium Enterprise)				
	(1)	(2)	(3)	(4)	(5)
INST	3.023 (0.560)	1.064 (0.300)	-0.156 (-0.050)	-0.137 (-0.040)	-0.545 (-0.180)
Sargan test (p-value)	0.384	0.679	0.474	0.317	0.726
Observations	57	55	57	55	55
	Panel B: First-Stage Regression (Dependent Variable: Measure of Institutional Quality)				
UK legal origin	0.476 (1.350)	0.292 (0.910)	0.890*** (2.960)		0.662** (2.350)
French legal origin	0.134 (0.350)	-0.045 (-0.130)	0.479 (1.490)		0.302 (1.030)
German legal origin	1.102** (2.010)	0.729 (1.490)	0.808* (1.780)		0.584 (1.420)
Scandinavian legal origin	0.838* (1.640)	0.481 (1.000)	0.526 (1.240)		0.361 (0.890)
AGYEAR		-0.0001*** (-3.300)		-0.0001*** (-4.140)	-0.0001*** (-3.110)
ABSLATIT			0.040*** (4.790)	0.031*** (4.370)	0.034*** (4.350)
F-test	2.45	5.29	7.52	22.56	9.44
Observations	57	55	57	55	55
R^2	0.629	0.729	0.756	0.779	0.814

Notes: this table presents instrumental variable (IV) regression estimates for the cross-section of countries. This table differs from table 14 in that it excludes the log of income per capita (LGDP), its quadratic form ($LGDP^2$) and the average years of schooling (SCH) from the explanatory variable list. Panel A reports the second-stage regression estimates from IV regressions with the relative size of the small enterprise sector versus the large enterprise sector (SEMSELE) as the dependent variable. "Sargan test" is the test of over-identification restrictions and is used to test the validity of the instruments. The numbers in parentheses are the z-statistics. Panel B presents the first-stage regression estimates with the measure of institutional quality (INST) as the dependent variable. "F-test" is the test for excluded instruments. The numbers in parentheses are the t-statistics. Conditioning variables used in both first-stage and second-stage regressions include log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), air distant (AIRDIST), employment laws index (EMPLAW) and civil rights index (CIVILR). We also include the number of observation and R^2 value for regressions in both first and second stages. Column 1 uses different legal origins -- UK legal origin, French legal origin, German legal origin and Scandinavian legal origin -- as instruments; Column 2 uses legal origins and agricultural transition year (AGYEAR); Column 3 uses legal origins and absolute latitude (ABSLATIT); Column 4 uses agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) and Column 5 uses legal origins, agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) as instruments. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Table 5.16: Results from IV estimation of the relationship between institutional quality and the relative size of small enterprise

Panel A: Second-Stage Regression

(Dependent Variable: The Relative Size of the Small Enterprise Sector versus the Large Enterprise Sector, SEMSELE)

	(1)	(2)	(3)	(4)	(5)
INST	5.176 (0.730)	0.800 (0.180)	0.905 (0.190)	10.345 (0.490)	-0.688 (-0.170)
FINDEV	-2.826 (-0.490)	0.551 (0.110)	-1.784 (-0.320)	-27.998 (-0.520)	0.285 (0.050)
Sargan test (p-value)	0.256	0.510	0.334	-	0.586
Observations	57	55	57	55	55

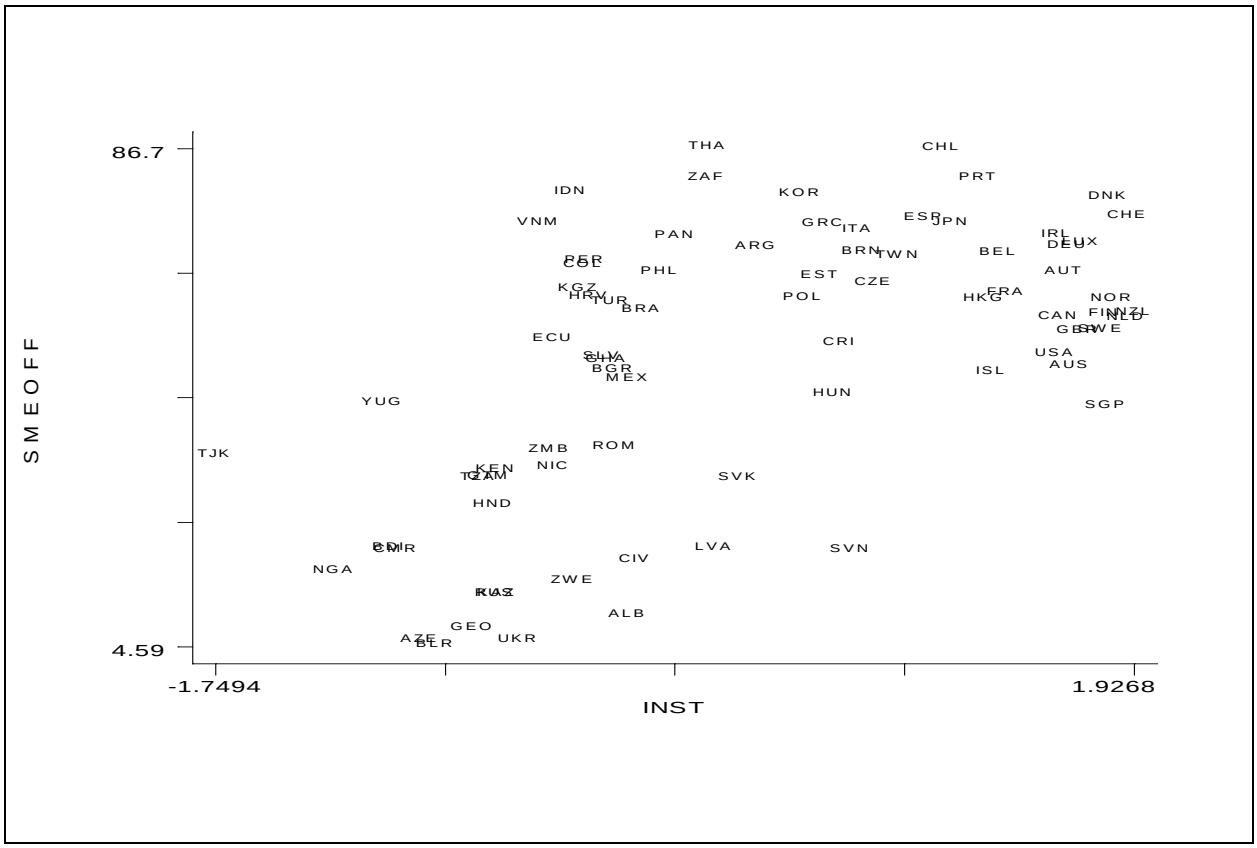
Panel B: First-Stage Regression

(Dependent Variable: Measure of Institutional Quality (INST) and the Level of Financial Development (FINDEV))

	INST	FINDEV	INST	FINDEV	INST	FINDEV	INST	FINDEV	INST	FINDEV
UK legal origin	0.475 (1.350)	0.669* (1.900)	0.292 (0.910)	0.670 (1.820)	0.890*** (2.960)	0.872** (2.450)			0.662** (2.350)	0.875** (2.320)
French legal origin	0.133 (0.350)	0.607 (1.590)	-0.045 (-0.130)	0.542 (1.400)	0.479 (1.490)	0.776** (2.040)			0.302 (0.103)	0.735* (1.870)
German legal origin	1.102* (2.010)	1.974*** (3.600)	0.729 (1.490)	1.864*** (3.310)	0.807* (1.780)	1.831*** (3.410)			0.584 (1.420)	1.785*** (3.240)
Scandinavian legal origin	0.837 (1.640)	0.457 (0.890)	0.481 (1.000)	0.508 (0.920)	0.525 (1.240)	0.305 (0.610)			0.361 (0.890)	0.442 (0.820)
AGYEAR			-0.0001*** (-3.300)	-0.00002 (-0.440)			-0.0001*** (-4.140)	-0.00003 (-0.670)	-0.0001** (-3.110)	-0.000 (-0.130)
ABSLATIT					0.040*** (4.790)	0.019* (1.980)	0.031*** (4.370)	0.016 (1.570)	0.034*** (4.350)	0.019* (1.800)
Partial R^2	0.179	0.247	0.386	0.248	0.461	0.307	0.501	0.071	0.580	0.302
Shea Partial R^2	0.149	0.205	0.347	0.222	0.322	0.215	0.056	0.008	0.428	0.224
Observation	57	57	55	55	57	57	55	55	55	55
Joint Sig Test (p-value)	0.603		0.862		0.713		0.660		0.887	

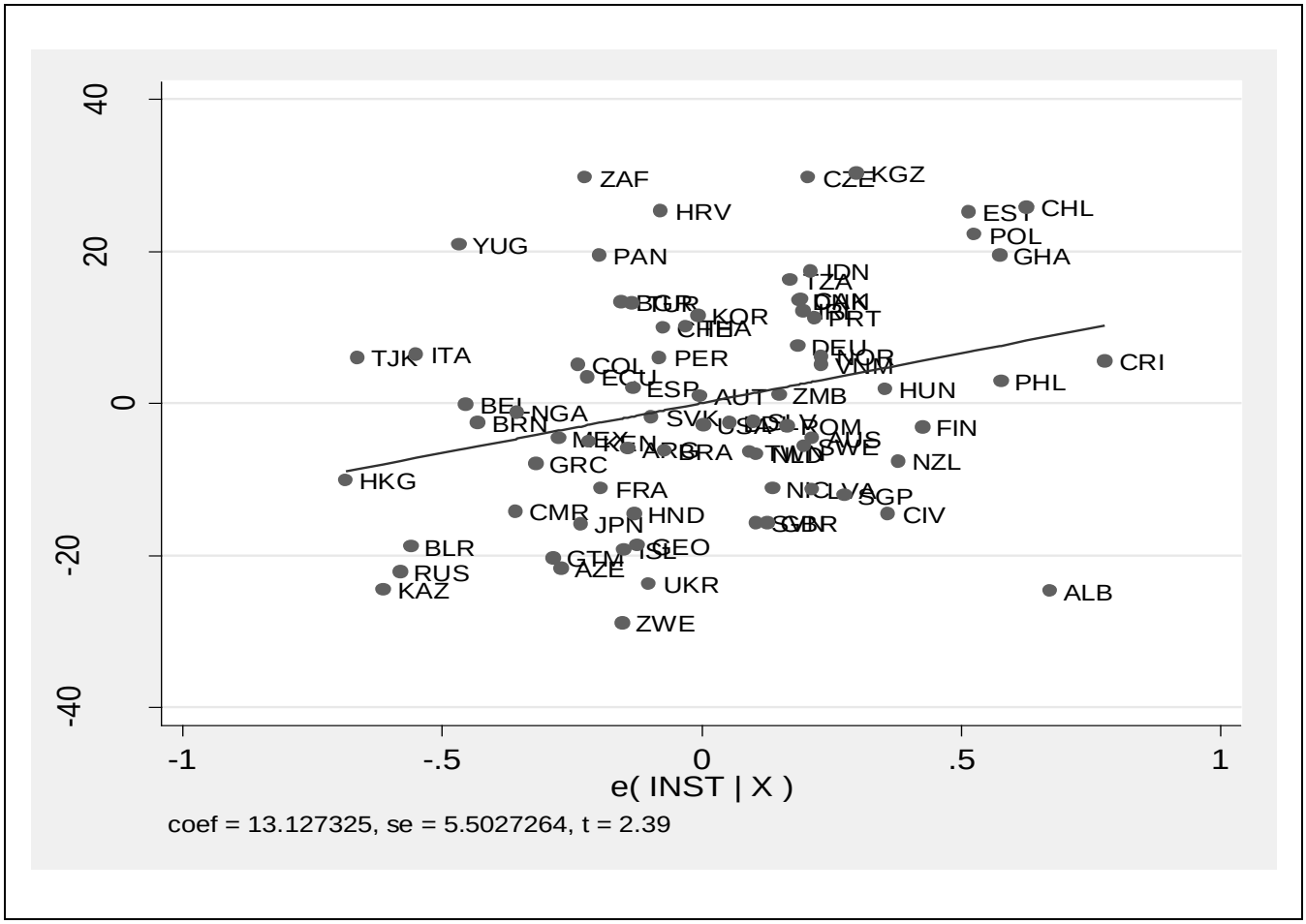
Notes: this table presents instrumental variable (IV) regression estimates for the cross-section of countries. Panel A reports the second-stage regression estimates from IV regressions with the relative size of the small enterprise sector versus the large enterprise sector (SEMSELE) as the dependent variable. "Sargan test" is the test of over-identification restrictions and is used to test the validity of the instruments. The numbers in parentheses are the z-statistics. Panel B presents the first-stage regression estimates with the measure of institutional quality (INST) as the dependent variable. "Partial R^2" is the R^2 of the first-stage regression with included instruments partialled-out. "Shea Partial R^2" is a partial R^2 that takes into account the intercorrelations among the instruments. "Joint Sig Test" is the test of joint significance of endogenous regressors. The numbers in parentheses are the t-statistics. Conditioning variables used in both first-stage and second-stage regressions include log of share of exports in GDP (LEXPGDP); continental dummy variables for Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM); share of service sector in GDP (SERVGDP); closest distance to a major port (AIRDIST); civil rights index (CIVILR) and employment law index (EMPLAW). We also include the number of observation and R^2 value for regressions in both first and second stage. Column 1 uses different legal origins -- UK legal origin, French legal origin, German legal origin and Scandinavian legal origin -- as instruments; Column 2 uses legal origins and agricultural transition year (AGYEAR); Column 3 uses legal origins and absolute latitude (ABSLATIT); Column 4 uses agricultural transition year (AGYEAR) and absolute latitude (ABSLATIT) and Column 5 uses legal origins agricultural transition year (AGYEAR), and absolute latitude (ABSLATIT) as instruments. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Figure 5.1: Correlation between Institutional Quality Measure (INST) and the SME Sector Share (SMEOFF)



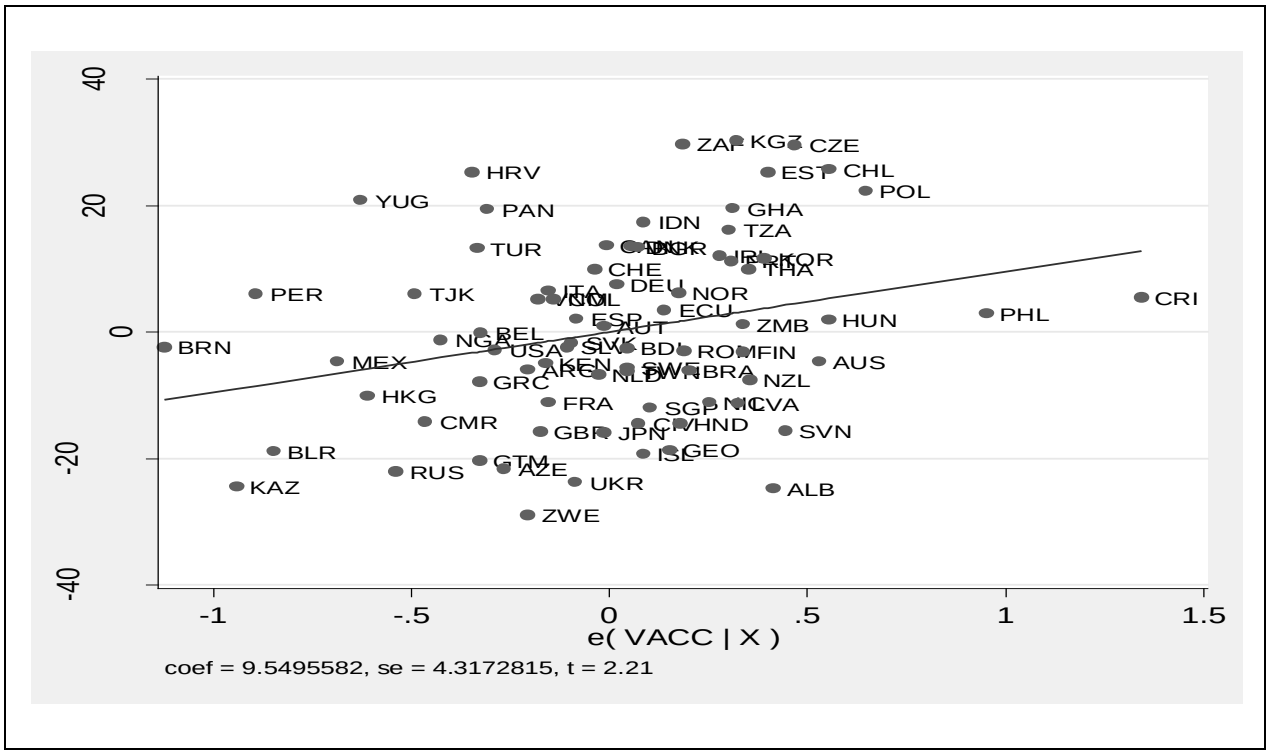
Notes: This figure displays simple correlation between the aggregate measure of institutional quality and the share of the SME sector

Figure 5.2: Partial Relationship between Institutional Quality Measure (INST) and the SME Sector Share (SMEOFF)



Notes: This figure shows the partial relationship between institutional quality measure and the SME sector share conditioning on other economic, social and geographical factors. Specifically, the list of conditioning variables includes log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average years of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), share of service sector in GDP (SERVGDP), air distant (AIRDIST), civil rights index (CIVILR), employment laws index (EMPLAW), and measure of financial development (FINDEV) .

Figure 5.3: Partial Relationship between Each Component of the Institutional Measure And the SME Sector Share (SMEOFF)



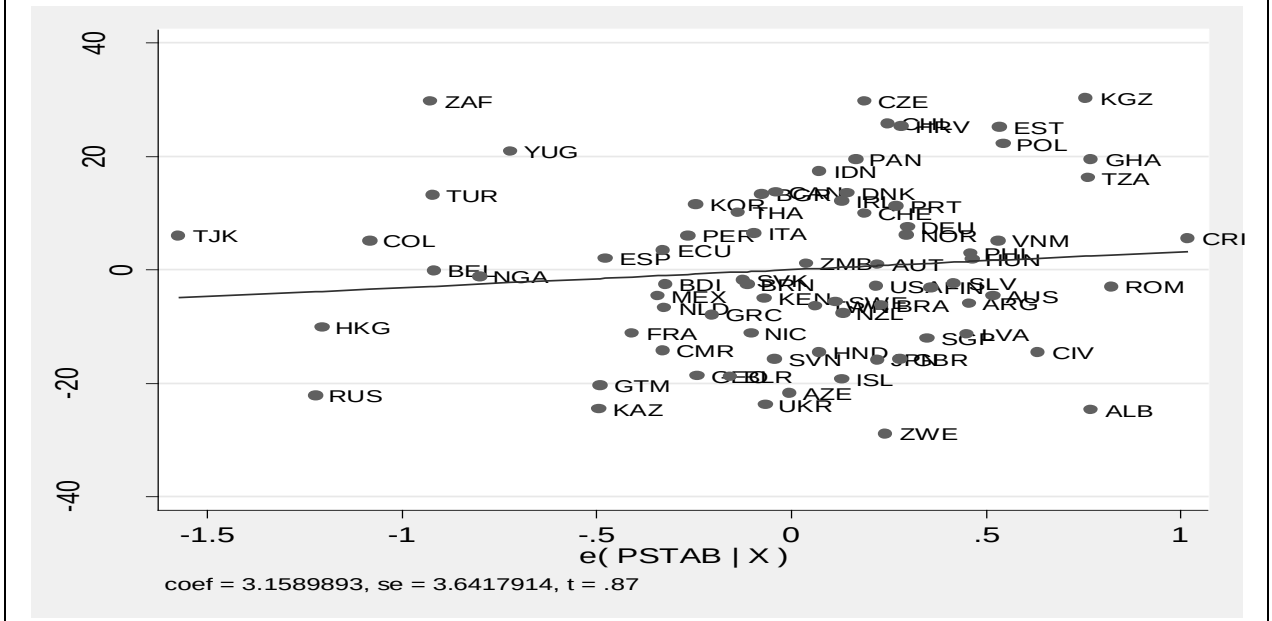
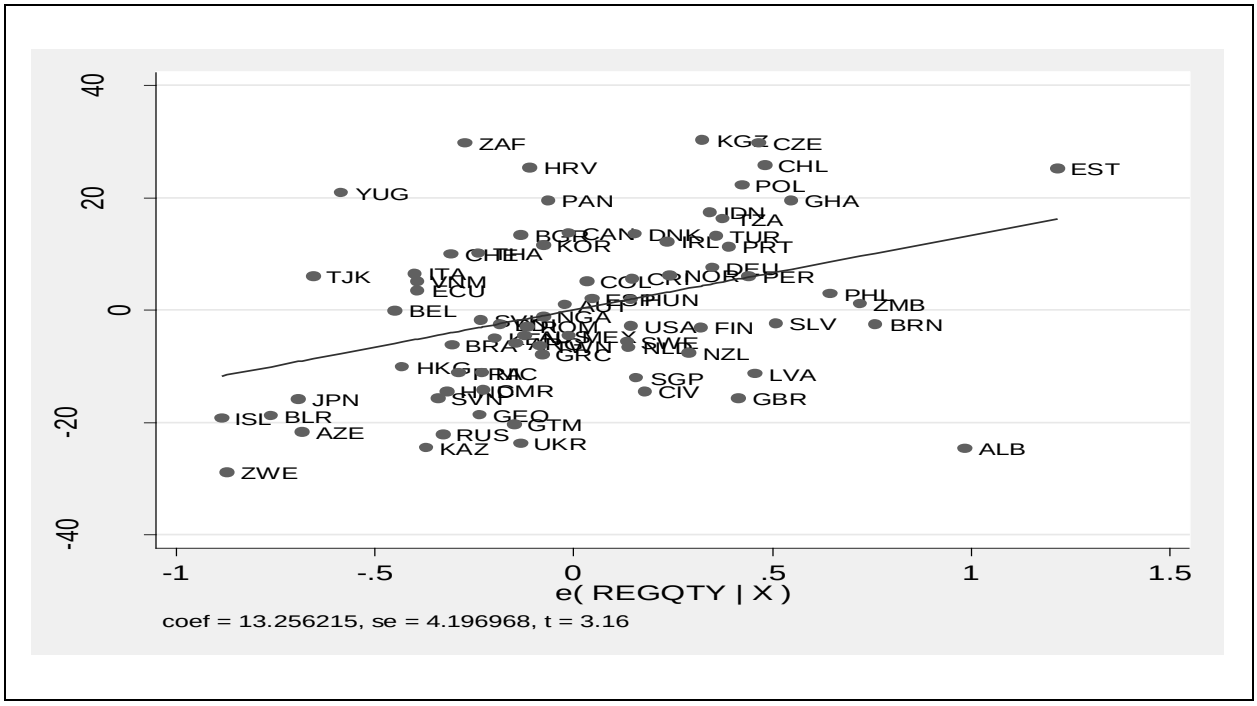


Figure 5.3 (Continue):



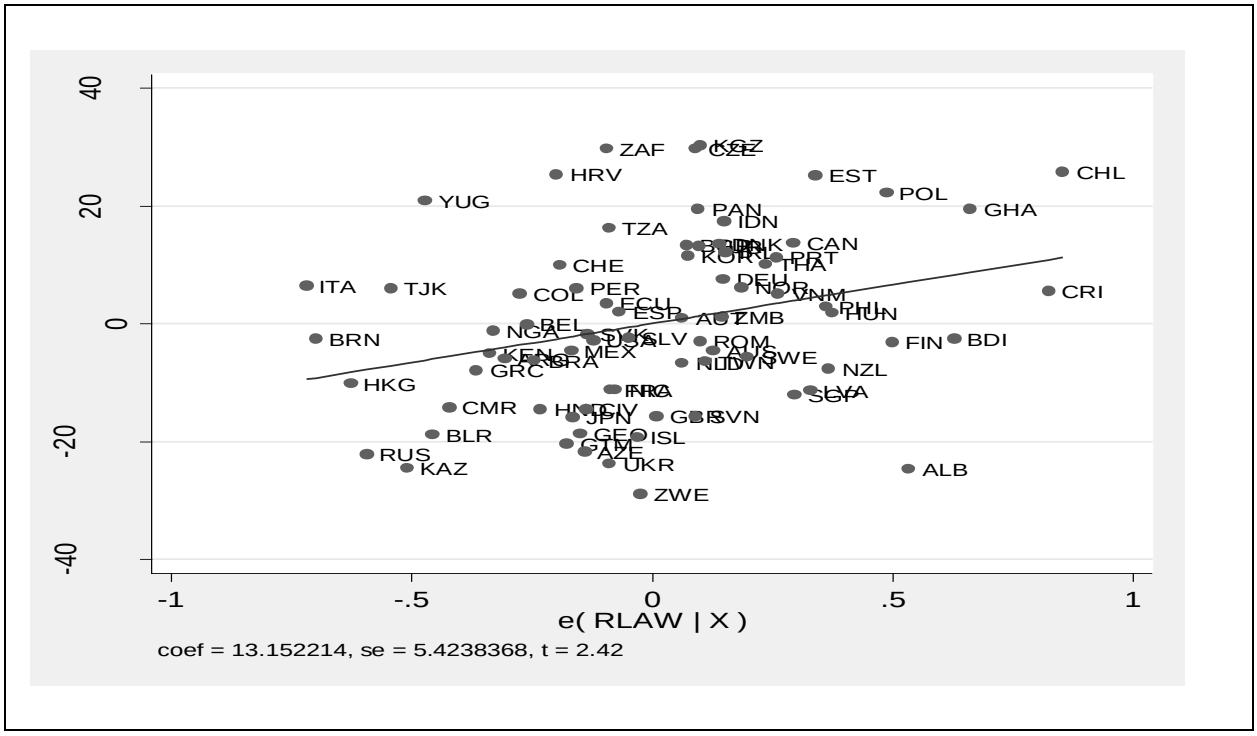
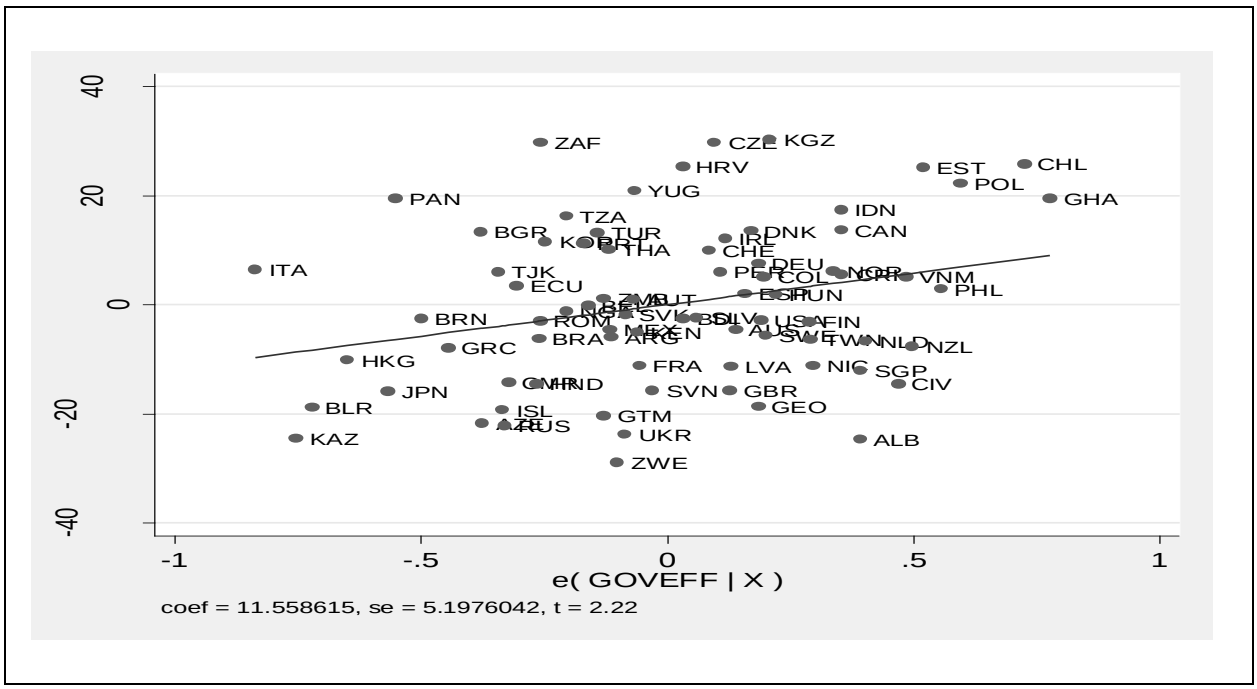
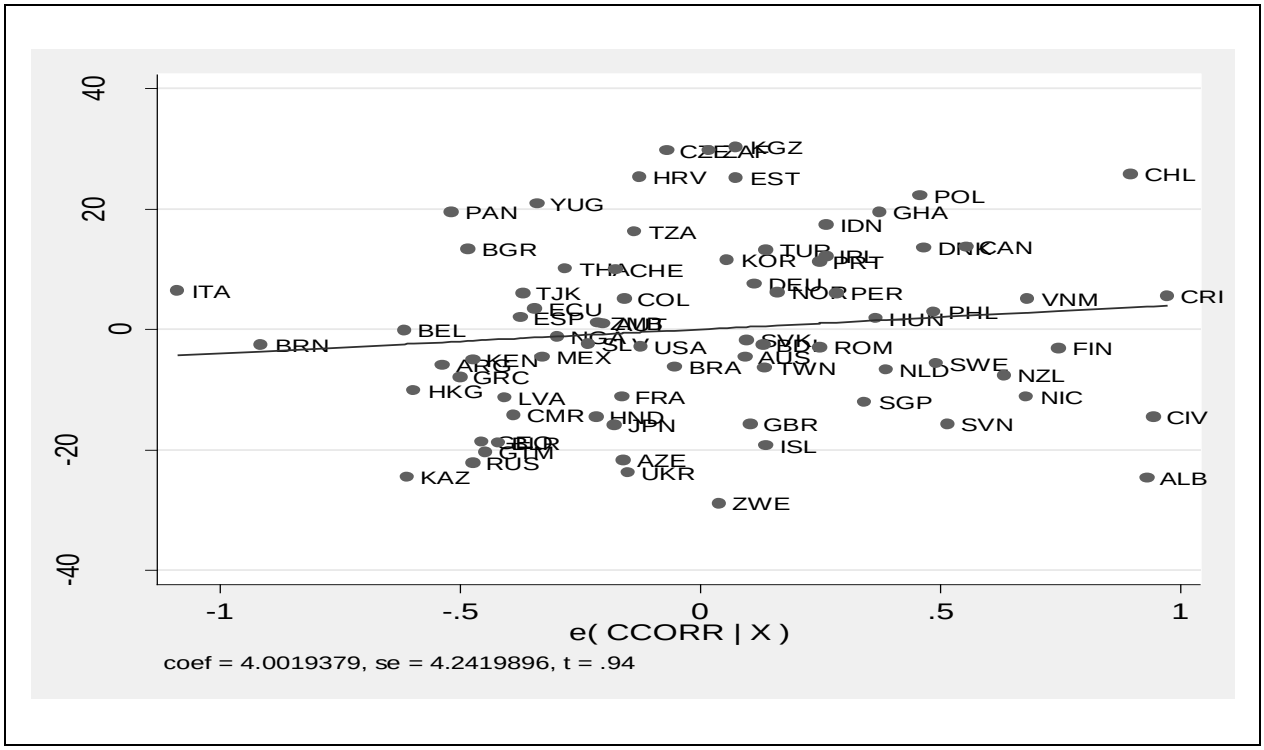


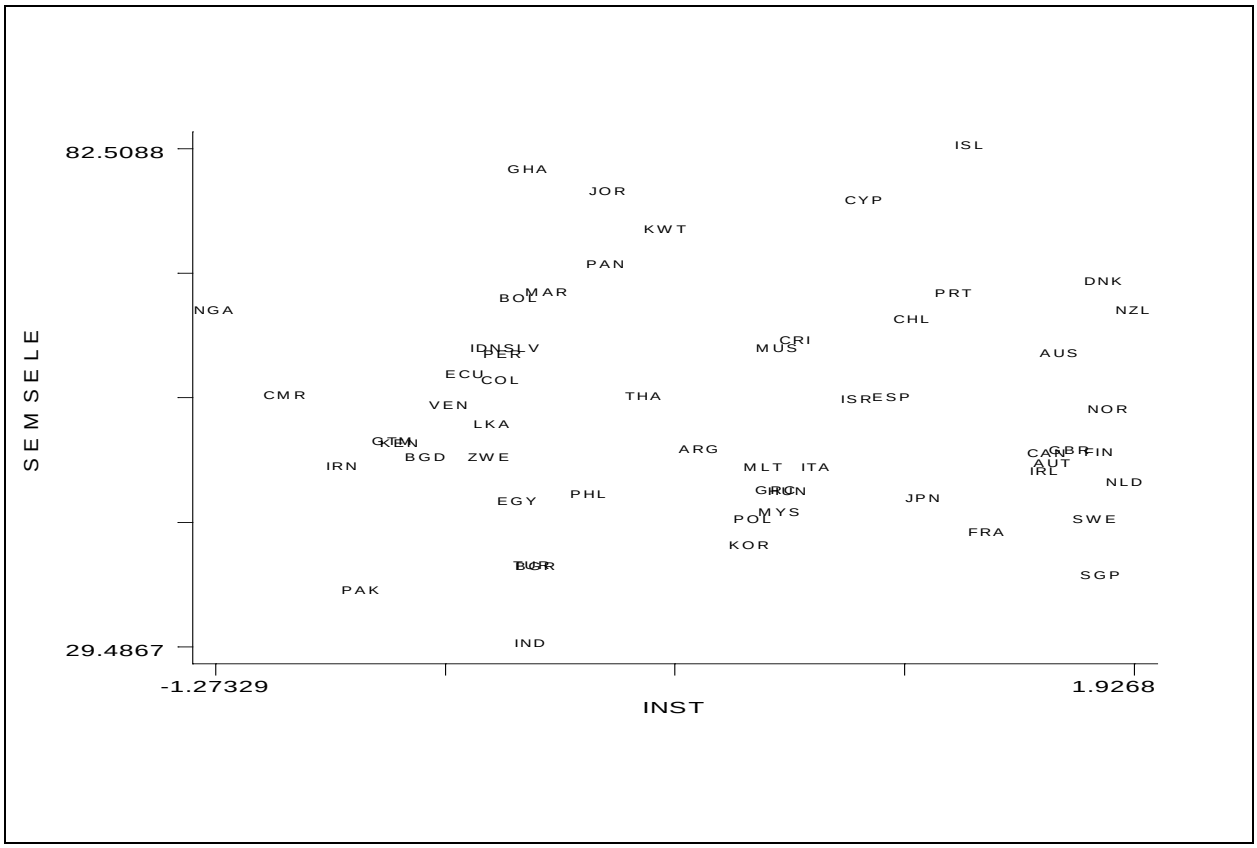
Figure 5.3 (Continue):





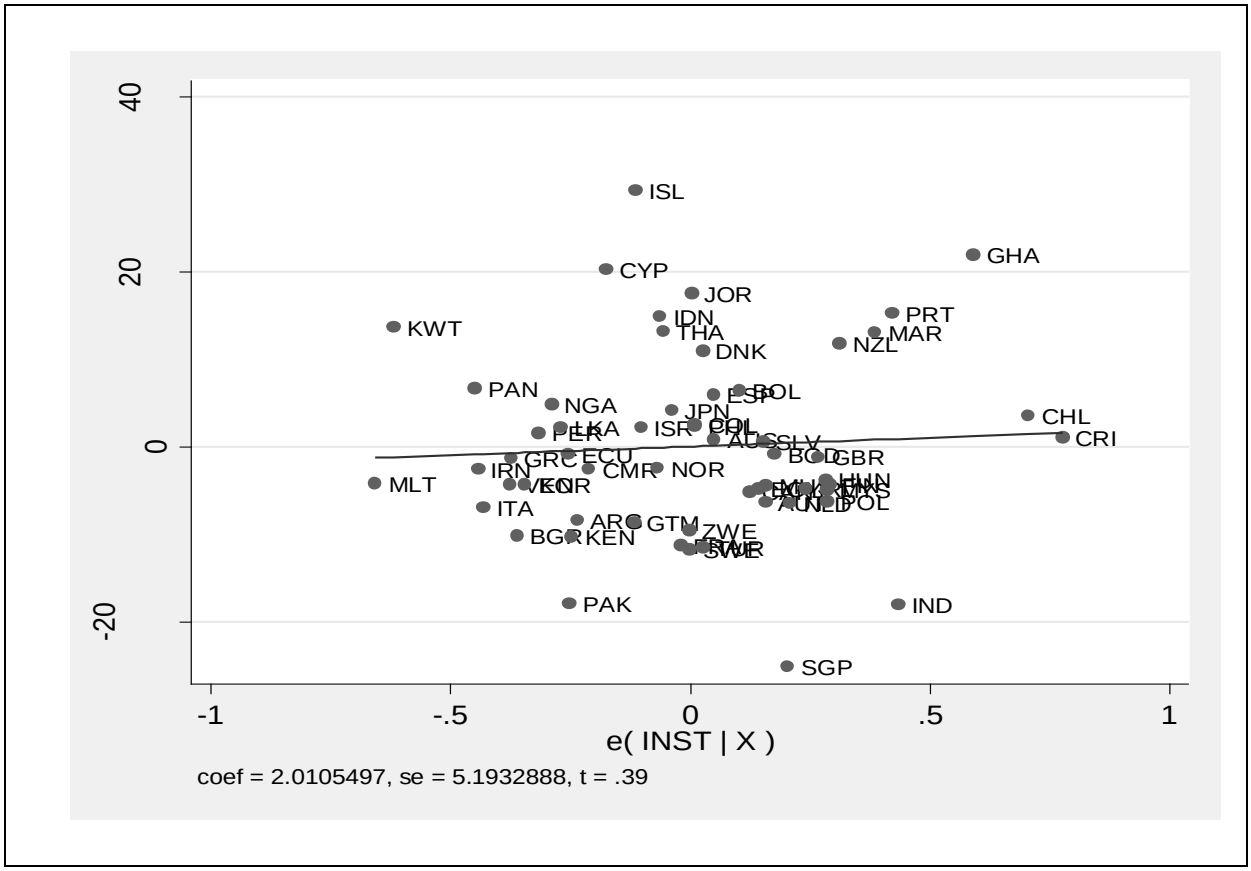
Notes: These figures show the partial relationships between different components of institutional quality measure and the SME sector share conditioning on other economic, social and geographical factors. Specifically, the list of conditioning variables includes log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average years of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), share of service sector in GDP (SERVGDP), air distant (AIRDIST), civil rights index (CIVILR), employment laws index (EMPLAW), and measure of financial development (FINDEV) .

Figure 5.4: Correlation between Institutional Quality Measure and the Relative Share of the Small Enterprise Sector (SEMSELE)



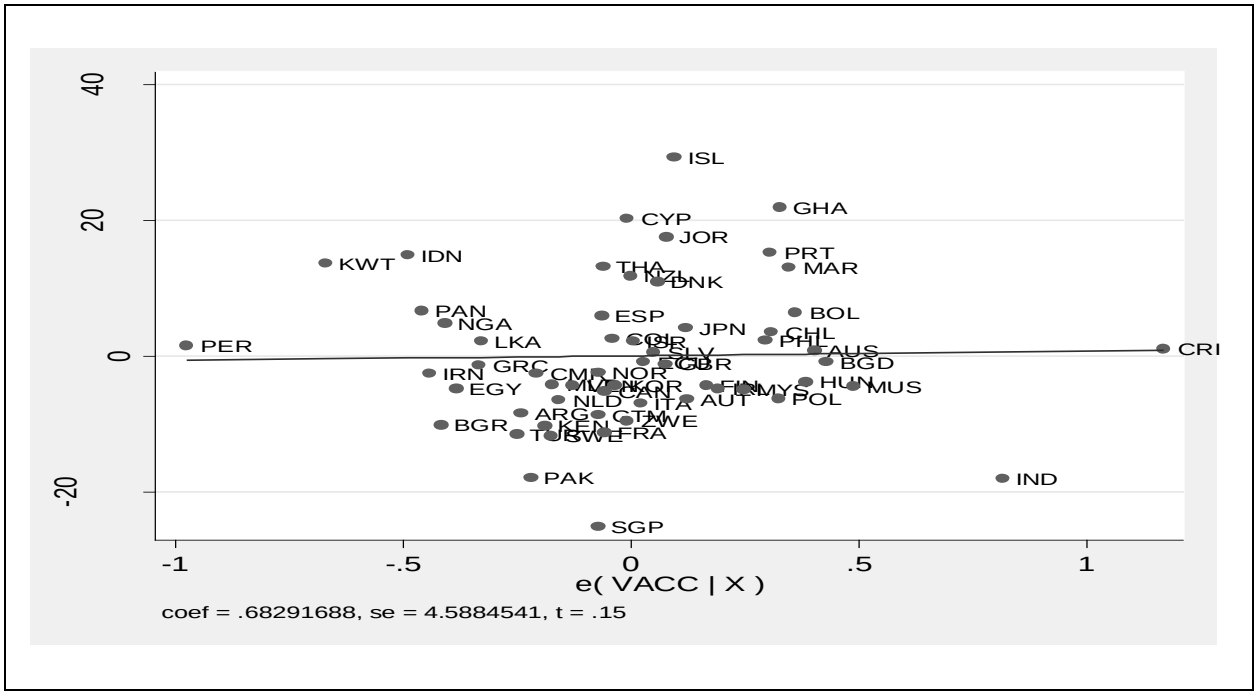
Notes: This figure displays simple correlation between the aggregate measure of institutional quality and the relative share of the small enterprise sector versus the large enterprise sector.

Figure 5.5: Partial Relationship between Institutional Quality Measure (INST) and the Relative Size of the Small Enterprise Sector versus the Large Enterprise Sector (SEMSELE).



Notes: This figure shows the partial relationship between institutional quality measure and the relative size of the small enterprise sector versus the large enterprise sector conditioning on other economic, social and geographical factors. Specifically, the list of conditioning variables includes log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average years of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), share of service sector in GDP (SERVGDP), air distant (AIRDIST), civil rights index (CIVILR), employment laws index (EMPLAW), and measure of financial development (FINDEV) .

Figure 5.6: Partial Relationship between Each Component of the Institutional Measure and the Relative Size of the Small Enterprise Sector Versus the Large Enterprise Sector.



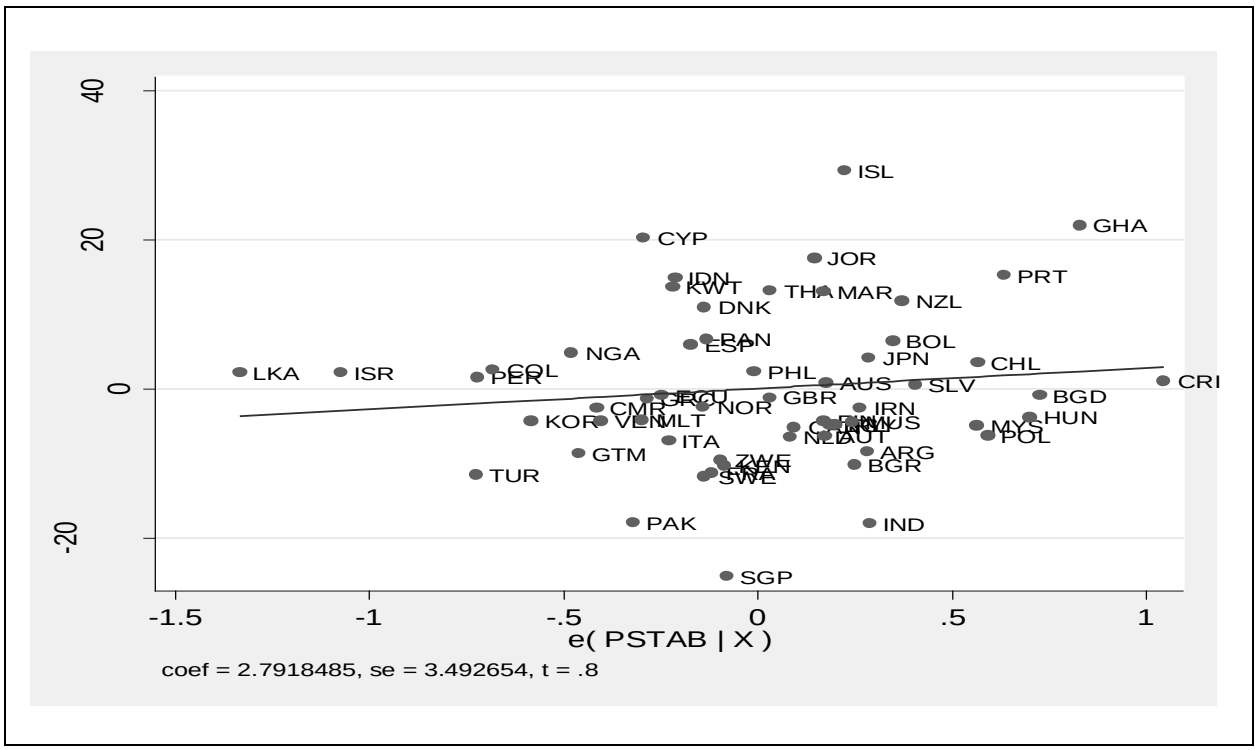
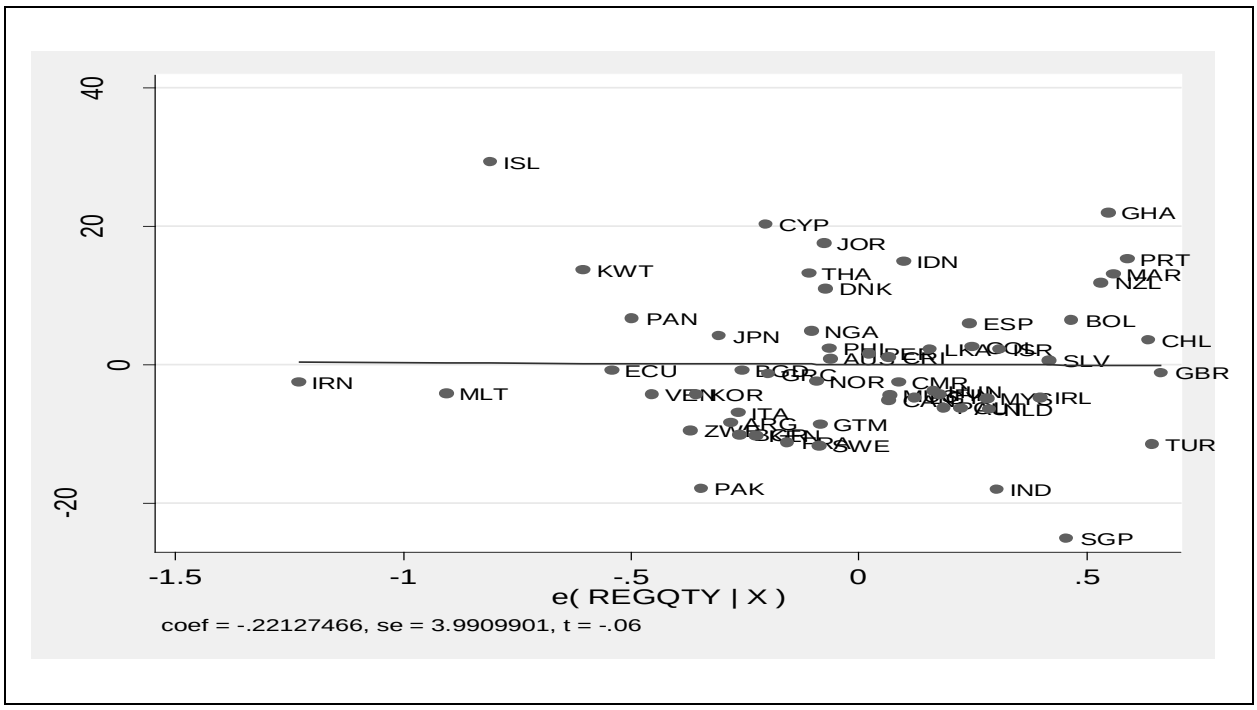


Figure 5.6 (Continue):



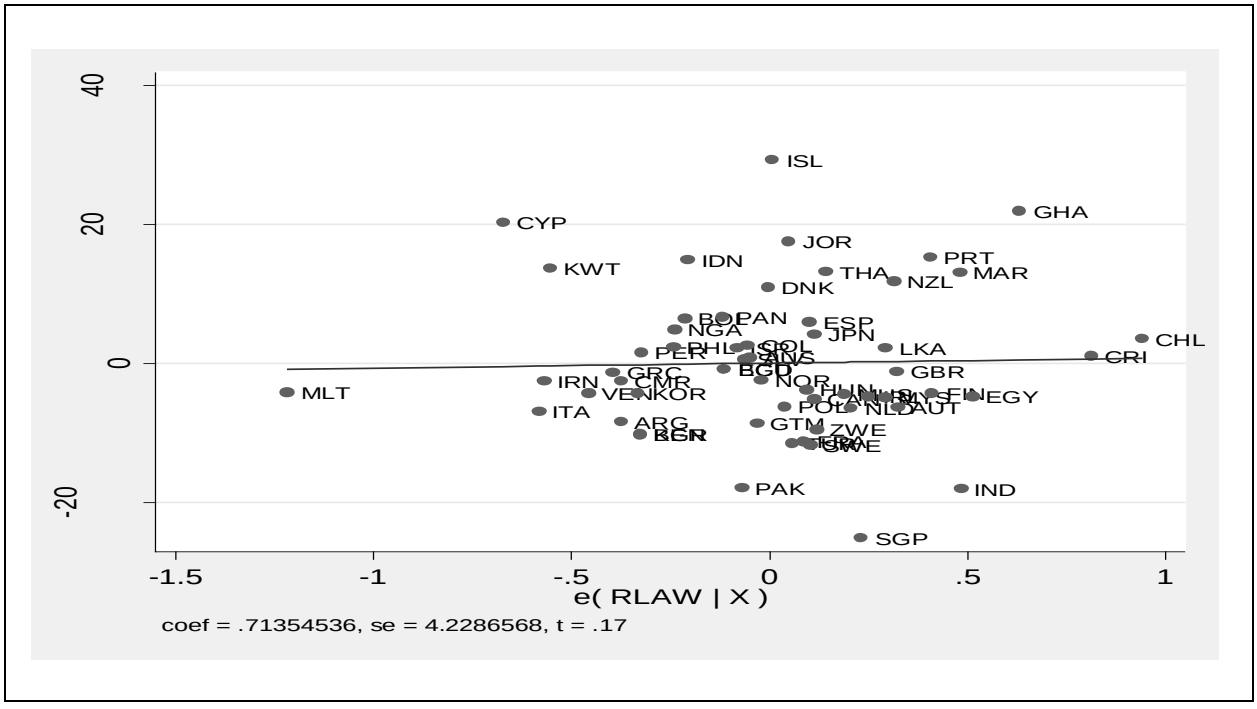
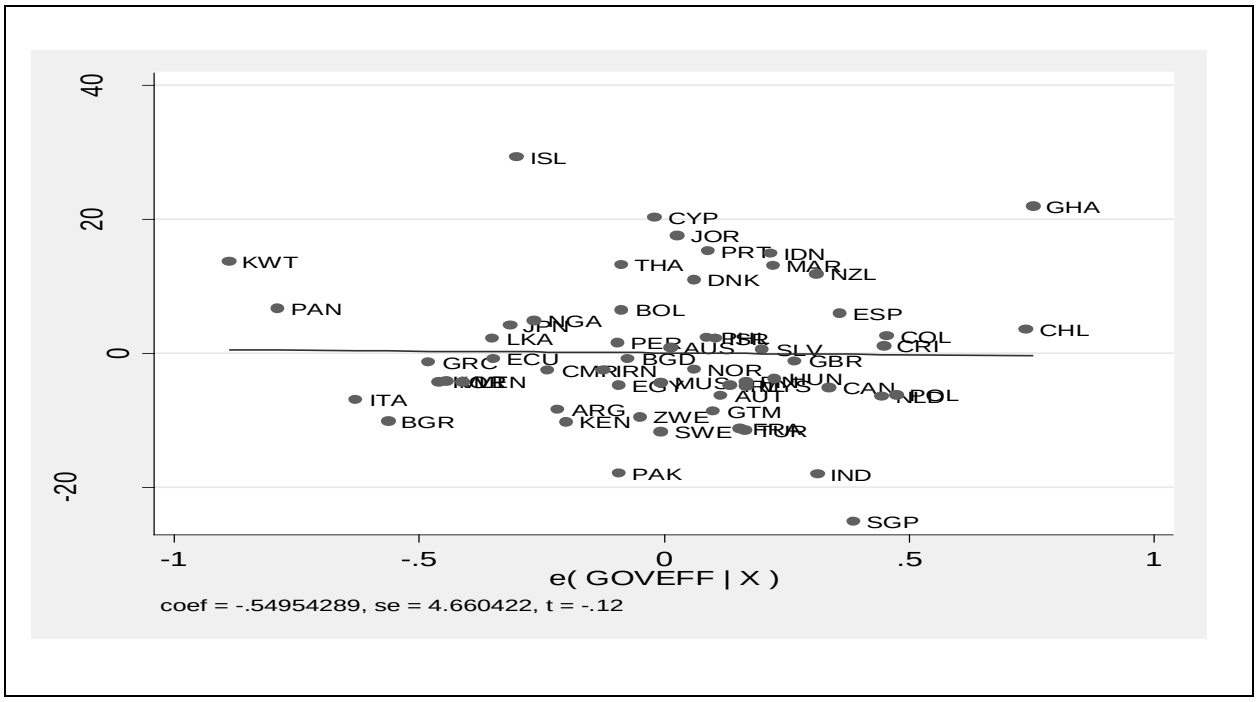
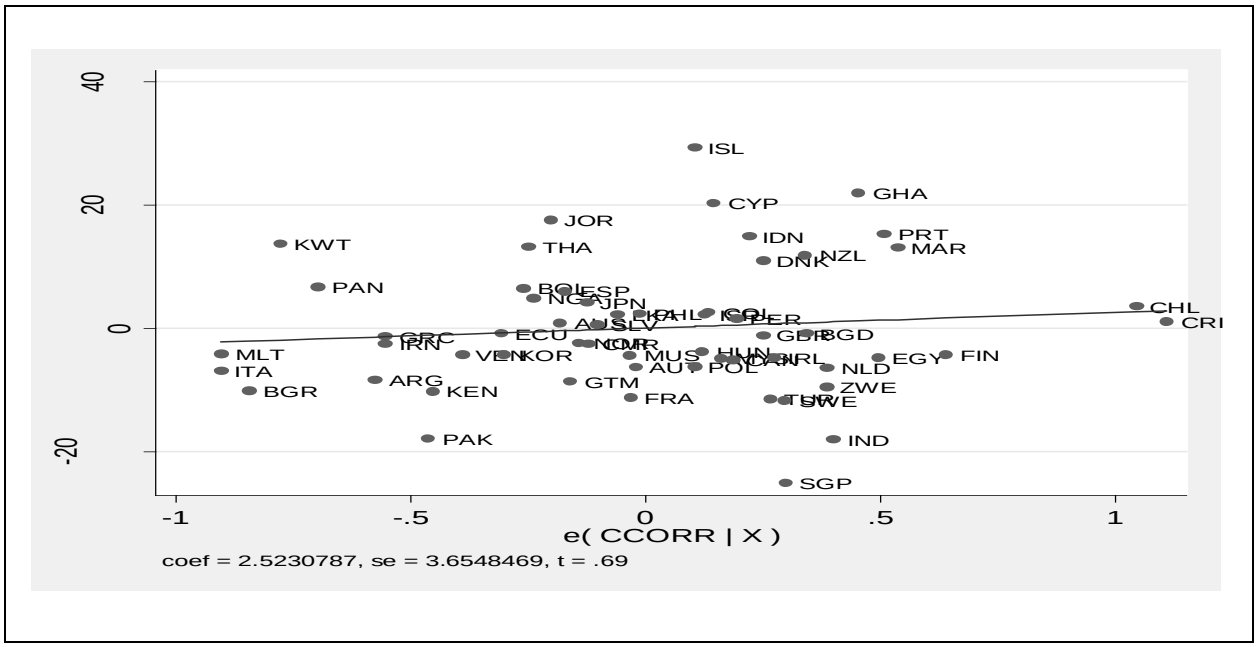


Figure 5.6 (Continue):





Notes: These figures show the partial relationship between different components of institutional quality measure and the relative size of the small enterprise sector versus the large enterprise sector conditioning on other economic, social and geographical factors. Specifically, the list of conditioning variables includes log of income per capita (LGDP) and its quadratic form ($LGDP^2$), average years of schooling (SCH), log of share of export in GDP (LEXPGDP), continental dummy variables for East Asia (ESEASIA), Europe (EU), Sub-Saharan Africa (SAFRI) and Latin America (LATAM), share of service sector in GDP (SERVGDP), air distant (AIRDIST), civil rights index (CIVILR), employment laws index (EMPLAW), and measure of financial development (FINDEV) .

Chapter 6
The Determinants of Structural Integration

Since Arthur Lewis first proposed it in 1954, the dualistic model of economic development and structural transformation of an economy has been a subject of interest in the study and analysis of the growth process, particularly for less developed countries. The central feature of dualism is the coexistence of a large agricultural sector with a small, active and dynamic industrial sector. Industry uses capital, and both sectors may undergo continuous technological change as they interact during the growth process. In this sense, dualism can be defined as a lack of intersectoral integration between these two sectors, namely the agricultural and the non-agricultural sectors.

One characteristic of a dualistic economy may be the existence of large wage and productivity gaps between the traditional and modern sectors implying a lack of integration between these sectors. In the classical dualistic model such as that of Lewis (1954) and Fei and Ranis (1961), the traditional agricultural sector is treated as a source of labour, which can be drawn to the capitalist sector at a constant real wage during the development process. Productivity in the traditional sector may be extremely low due to the abundance of labour with a fixed quantity of land, while the modern sector is characterized by higher productivity due to available advanced technology employed in that sector. Likewise, a wage differential between sectors is assumed to be an inherent characteristic of the economy in the sense that the wage in the modern sector needs to be always higher than that in the traditional sector in order to induce traditional-modern sector labour migration required for the growth process.

There is virtually no previous work, however, on the sources of cross-country differences in the extent of dualism. This chapter thus intends to fill this gap by using cross-country regressions to determine possible linkages between the degree of economic dualism, as proxied by the relative labour productivity of agriculture versus other sectors,[75]

[75] We follow Bourguignon and Morrisson (1998) who are among the few writers who have used relative labor productivity in agriculture (RLP), which they define as the ratio of value added per worker in the non-agricultural sector to that in agriculture, as a proxy for the degree of dualism in an economy.

and a few variables of interest including financial, political, macroeconomic, as well as a number of geographical variables. The results will show some of the factors that explain cross-country variation in the degree of dualism. However, it is important to note that they will not tell us anything about the direction of causality of these relationships.

Bayesian Model Averaging is used to investigate a cross-country dataset in order to determine the selection of models while minimizing the exposure to some of the common problems faced by conventional empirical work. For instance, the Bayesian approach acknowledges uncertainty about the appropriate regression specification, avoids some of the arbitrary choices that are often associated with cross-country empirical work, and provides an index of the weight of evidence in favour of specific models (Malik and Temple, 2004).

The estimation of relative labour productivity, the indicator of the degree of dualism, will be discussed in a later section of this chapter. Briefly, it equals the ratio of average labour productivity (or output per worker) in the non-agricultural sector to that in agricultural sector.

This study uses many explanatory variables which may have a possible relationship with dualism. Among them is a measure of the level of economic development as proxied by GDP per capita measured in 1995. This variable is picked because of its possible relationship with the degree of dualism as proposed in a few previous studies such as those by Kuznets (1971), Chenery and Syrquin (1975), and Ahluwalia (1976). We also consider a number of financial variables as, according to the classical dualistic model of development, financial capital plays a vital role in promoting growth and reducing the degree of intersectoral distortion. Other variables to look at include human capital, geography, and political system variables.

Similar to previous work we find an inverted U-shape relationship between relative labour productivity and the level of income. During the early stage of development, intersectoral wage and productivity gaps tend to increase because of the asymmetric distribution of scarce capital in favour of the modern sector, while the traditional sector is left with excessive labour. This gap then gradually narrows down as relative productivity in agricultural rises due to the modernization of agricultural technology and as the surplus of agricultural labour is absorbed by the rest of the economy.

Our results show a less clear relationship between financial variables and the degree of dualism. On the one hand, the analysis indicates a negative relationship between the initial level of capital accumulation – as proxied by percentage of gross domestic saving in gross domestic product – and the relative productivity of labour. This supports the proposition that capital accumulation plays a vital role in promoting growth and structural integration of the economy. On the other hand, the relationship between the level of financial development and the degree of dualism is less clear. Despite the vital role that finance may play in promoting growth and intersectoral integration, our analysis gives a rather ambiguous picture of this relationship.

We also find that landlocked countries tend to be associated with a more dualistic structure of the economy. This might be due to the fact that, because of their lack of access to shipping routes for international trading, these countries tend to concentrate their activities in agriculture and in exporting primary products, rather than promoting a modern sector that usually involves producing high value added goods for export and trade. We also find positive relationships between the degree of dualism and a few other geography variables, including tropical location and a dummy for African countries.

A more politically open society is also found to be associated with a higher degree of dualism which may be due to a higher degree of intersectoral wage distortion. One possible hypothesis is that urban workers in these countries usually wield more bargaining power through various channels such as labour unions and thus are more able to secure higher wages than if wages were set by the market.

Furthermore, we find the initial level of human capital does a better job than the current level in explaining the cross-country variation in dualism. However, we find the roles of the population growth rate and a few other demographic variables to be non-robust.

The next section of this paper will present discussions of possible relationships between relative labour productivity and various factors. Section 6.2 will be a survey of the literature on the subject of dualism. In section 6.3 we will describe the sources of our data and also the variables used in our study. Section 6.4 will present the empirical results. Finally, the last section will be the conclusion.

6.1 Possible Factors Relating to Dualism

In this section we will present discussions on the potential relationships between a number of factors and the degree of dualism, as reflected in the value of relative labour productivity. We will focus on a number of variables including the level of economic development and of financial development, human capital, political system, geographical factors and a few other variables.

In the classical dualistic model of development proposed by Lewis (1954) and Fei and Ranis (1961), during the early phases of economic growth, wage and productivity gaps between traditional and modern sectors tend to increase as growth of capital formation goes to improve only the wage and productivity of workers in the modern sector while wages and productivity of labour in the traditional/agricultural sector remain unchanged at a very low, subsistence level. Only at the later stage of development does this gap begin to narrow as capital accumulation in the modern sector works to eliminate labour surplus and disguised unemployment in the traditional sector.[76]

Such a proposition is supported by other writers such as Kuznets (1971), Chenery and Syrquin (1975) and Ahluwalia (1976) who suggest that there exists an inverted U-shape relationship between development and the extent of dualism. They argue that during the development process, the relative labour productivity of the primary sector versus other sectors falls initially during the early phase of transformation as scarce capital and other resources are concentrated in the modern sector to the detriment of productivity in the traditional agricultural sector. This works to widen the intersectoral productivity gap during the early stage of development. This gap then gradually narrows as relative productivity in agriculture rises due to the modernization of agricultural technology and as the surplus agricultural labour is absorbed by the rest of the economy. The productivity gap is then largest at the middle income range.

There have been a number of studies on the relationship between the level of financial development and economic growth. A well functioning financial system is said to be very

[76] Some writers such as Fei and Ranis (1961), Reynolds (1969) and Jorgenson (1967) refer to this as the commercialization point where the marginal productivity of labor, and with this the earnings of labor, are equalized in the two sectors in the sense that labor becomes a scarce resource for which agriculture and industry must compete.

important in promoting economic growth via its role in mitigating risks, facilitating the exchange of goods and services, mobilizing savings and allocating resources contributing to capital accumulation and technological innovation which is beneficial for growth. King and Levine (1993) and Levine (1997) find a strong positive association between the level of financial development and the rate of economic growth. Furthermore, Levine (1997) suggests that insufficient financial development has sometimes created a poverty trap and thus becomes a severe obstacle to growth even when a country has established other conditions for sustained economic development.

Joseph Schumpeter (1911) argued that banks play a pivotal role in economic development because they choose which firms get to use society's savings. According to this view, the banking sector alters the path of economic progress by affecting the allocation of savings and not necessarily by altering the saving rate. Beck, Levine and Loayza (2000a, 2000b) use cross-section and panel data for a sample of 63 countries over the period 1960-95 and have confirmed this view. They found that banking sector development exerts a strong and robust causal impact on real per capita growth and per capita productivity growth. However, they do not find a robust relationship between banking sector development and either physical capital accumulation or private saving rates. Thus, their result conforms to Schumpeter's view that banks affect economic development primarily by influencing total factor productivity growth.

A possible channel through which the level of financial development and dualism are associated with each other is the level of efficiency of resource allocation provided by the financial system. One of the characteristics of an underdeveloped financial system is the existence of a high degree of financial dualism. Financial dualism refers to a lack of integration between capital markets operating in the traditional sector and the modern sector. Under financial dualism, larger size firms in the modern sector have better access to funds on easier terms than small economic units in the traditional sector. Firms in the modern sector can borrow at relatively low interest rates from the 'organized' capital market while peasant farmers or handicraft workers in the traditional sector do not have access to such institutional sources of credit and thus must borrow at high interest rates from the 'unorganized' capital market.

In addition, the gap in interest rates tends to be larger the more underdeveloped are domestic credit markets and financial institutions in the country. Financial dualism may cause an excessive allocation of capital to the modern sector relative to the traditional sector (Meier, 1989). It follows that such a distortion in credit allocation under the regime of financial dualism can fuel economic dualism as well, especially at the early stage of development.

McKinnon (1973) argues that the differential access to credit produced by financial repression fosters economic dualism in developing countries, defined as the coexistence of traditional techniques with low productivity and low income generation and modern techniques with high productivity and high income generation. He also suggests that providing low-cost credit to some and denying it to other entrepreneurs will result in investment inefficiency and income inequality. Cho (1984) shows that financial underdevelopment due to financial repression is likely to widen wage gaps between high skilled labour in the modern sector and low skilled labour in the traditional sector. Dualism, taking the form of inefficient small-scale direct investments on the one hand and excessively capital-intensive large-scale investments on the other hand, creates greater dispersion in wages. The reason is that capital intensive-production methods in the modern sector reduce the demand for unskilled labour, which in turn, causes the wages of unskilled labour to fall.

Krugman (1978), however, presents a slightly different conclusion. He shows that while financial repression does indeed cause income inequality and investment inefficiency, it does not necessarily cause economic dualism. He suggests that even if all individuals had identical tastes, endowments, and access to both technologies, some would still choose traditional technology and others the modern technology. Indifference in this choice springs from the fact that both yield exactly the same levels of consumption when interest rates are market-determined.

One possible remedy to the problem of financial dualism is suggested by Myint (1971), who argues that in order to reduce financial dualism between the organized and unorganized markets for credit, there would have to be more effective links facilitating the mobility of funds between the organized and unorganized credit markets. In this sense we can expect that financial development which helps bring about such a closer link, and thus a more efficient credit allocation, would not only promote faster economic growth, but would

also reduce both the degree of financial dualism and economic dualism. This implies that a higher level of financial development may be associated with a lower degree of economic dualism.

Human capital is another mechanism that might be related to the degree of dualism during the development process. According to Barro (1991), human capital plays a very important role in economic growth by allowing a country the ability to command a more productive labour force and to adopt better technologies, which in turn leads to a higher rate of growth and development. Also as Ahluwalia (1976) points out, there is the argument that skill intensive development patterns are less prone to be concentrated in one sector than capital intensive patterns. This is because of the peculiar characteristic of human capital – unlike physical capital – that expansion in the stock of human capital in the economy may involve dispersion across a wider population. In addition, human capital in the traditional sector is a very important factor which enables the adoption of new techniques and technologies leading to an increase in productivity in this sector and thus reducing the intersectoral productivity gap.

Another possible channel in which human capital relates to dualism is through its potential role in inducing intersectoral shifts of labour. Education seems to be an important factor influencing rural-urban migration. Numerous studies of migration in diverse countries have documented the positive relationship between the educational attainment of an individual and his or her propensity to migrate from rural to urban areas. Individuals with higher levels of education face wider urban-rural real-income differentials and higher probabilities of obtaining modern sector jobs than those with a lower level of education (Todaro and Smith, 2003). Moreover, Chenery and Syrquin (1975) show that, with the exception of a few countries, after a certain income level rural-urban migration has reduced the labour productivity differential substantially and that the share of industry and services in the labour force is much larger and becomes closer to their shares in production. In these ways, a higher level of human capital would in the long run work to reduce intersectoral productivity gaps and thus the extent of dualism.

The degree of political openness can be said to have an indirect relationship with dualism through its effect on the intersectoral wage differential. Political freedom enables people to participate freely in the political process. Usually in a freer society, such orga-

nizations as labour unions have more bargaining power and are potentially more capable of keeping the wages of union members somewhat higher than those set by the market. Thus, some forms of wage distortion may be more likely to exist in a freer society than in closed ones and there are cases where democratic governments adopt potentially counter-productive labour market policies. However, as Banerji and Ghanem (1997) suggest, this view does not always hold true. There is a view that the probability of governments passing inefficient labour legislation to benefit insiders is higher under authoritarian than under democratic regimes.

The role of geography in explaining dualism should also be taken into consideration because of its relationship with sectoral development in the economy. Sachs (2003) emphasizes a strong relationship between underdevelopment and tropical location. Tropical countries face several obstacles to economic growth including remoteness from large markets, a higher incidence of disease, poor natural resource endowment, and climatic factors. Similarly, landlocked countries face obstacles when it comes to promoting trade because of high costs of engaging in such trade, especially the cost of transportation.

Such lack of access to external markets can reduce a country's ability to diversify its export production and become dependent on its natural resources, which can be detrimental to long term growth and development. Several studies such as that of Radelet and Sachs (1998) and Redding and Venables (2003) indicate that geographically remote countries will find it harder to develop non-primary exports, and especially manufacturing goods. The production of labour-intensive manufactured exports is often associated with a high import content and small profit margins, so natural barriers to trade can easily render such exports uncompetitive. Thus the cumulative effects of the remoteness of geographical location tends to inhibit the country's ability to develop a more dynamic sector, notably the manufacturing sector, while increasing its reliance on a more basic sector, such as a traditional /primary goods sector – which is one of the main problems of underdevelopment. In our analysis, we will examine the possible effects that geographical barriers to trade may have on the extent of dualism.

Finally, a high population growth rate can also be a factor contributing to dualism. Hollis Chenery (1979) suggests that one of the principal causes of a slow rise in average agricultural output per person employed in primary production during the development

process is the rapid growth rate of population. In countries without a surplus of labour and with fairly equal access of all sectors to capital and technical knowledge, this lag has not been observed. These studies, however, focus mainly on the relationship between population growth and the pace of development, while its effects on the degree of structural integration or dualism have hardly received any attention. One possible link between population growth and the degree of dualism lies in the difference in the rate of population growth between different sectors of the economy. At the early stage of development, population growth in the traditional, agricultural sector normally grows at a faster pace that in the modern, urban sector. This fact and the intersectoral technological asymmetry could give rise to a disequilibrium in wages and productivity between the traditional and modern sectors.

6.2 Previous Studies

In this section we will discuss the various schools of thought regarding the development process of a dual economy. We will first present the classical and neoclassical views, and their assumptions about the development process of a dualistic economy. Then we discuss some of the criticisms of these models.

The development process of a dualistic economy is normally considered to be driven by capital accumulation in the modern, capitalist sector of the economy which enables it to absorb excess labour in the traditional sector into the growth process. The classical dualistic models such as those of Lewis (1954) and Fei and Ranis (1961) present a detailed analysis of the intersetoral interaction between the traditional and the modern sectors which enables economic growth and development to take place.

One of the main assumptions of these models is the existence of a labour surplus in the traditional agricultural sector. Labour surplus is said to exist when the marginal productivity of labour in agriculture is extremely low, perhaps close to zero, and the removal of this surplus labour from agriculture will not reduce total output in that sector. Further, while productivity in the traditional sector may be extremely low due to the abundance of labour with a fixed quantity of land, the modern sector is characterized by higher productivity due

to advanced technology available to that sector. It follows that dualism can also be characterized by the gap between labour productivity in modern and traditional sectors. Thus, the more dualistic the economy is, the wider the spread between the labour productivity in modern and traditional sectors and the higher the value of relative labour productivity.

Another important assumption of the model relates to the wage differential between traditional and modern sectors. The level of the wage in the urban sector is assumed to be constant and determined as a given premium over a fixed average subsistence level of wages in the traditional agricultural sector. As Lewis puts it, “earnings in the subsistence sector sets a floor to wages in the capitalist sector, but in practice the wage has to be higher than this, and there is usually a gap of 30 percent or more between capitalist wages and subsistence earnings" (Lewis, 1954).

The neoclassical model of dualism, proposed by Jorgenson (1961), discusses the conditions for the industrial sector to emerge from the purely agrarian economy (the viability condition for take-off into sustained growth). According to this model, there exists a critical level of food production in the agricultural sector which agricultural average product needs to exceed, in order to create an agrarian food surplus and thus generate a sufficient supply of both labour and food to enable the development of the industrial sector to begin. For Jorgenson’s model, therefore, the necessity condition for sustained growth is a positive and growing agricultural surplus with or without the existence of disguised unemployment. Furthermore, the combination of a positive and growing agricultural surplus and a small positive initial capital endowment is sufficient for take-off into sustained capital accumulation and output increase (Jorgenson, 1961). In other words, a massive infusion of capital is not necessary for development leading to sustained growth.

Both the classical and neoclassical models of dualism place a strong emphasis on the vital role of capital accumulation in promoting economic growth and structural transformation. As mentioned earlier, capital accumulation is assumed to exist only in the modern sector and is the main factor allowing this sector to develop by drawing surplus labour from the traditional sector as needed to fuel the development process. Furthermore, Lewis argues that as more capital becomes available, more workers can be drawn into the capitalist sector from the subsistence sector, and their output per head rises as they move from one sector

to the other – i.e. labour productivity increases with a labour transfer from agriculture to the capitalist sector.

There are, however, problems with some of the underlying assumptions of these models. First, classical and neoclassical models of dualism assume that the output of the traditional sector is a function of land and labour alone; there is no accumulation of capital except in the form of land reclamation and technical progress in the agricultural sector is assumed to be exogenous and neutral. However, Dixit (1970) argues that assumptions of exogeneity and neutrality of technical progress are poor assumptions in the context of a labour surplus. His reasoning is that if there is surplus labour, a very important role of technical progress is that of rendering this labour productive. Technical progress in a real labour-surplus economy will clearly be induced by this necessity. In addition, Niho (1974) points out that while the assumption of no capital accumulation in agriculture may apply to contemporary developing countries, it has not always been the case as historical evidence show that the application of capital inputs (such as fertilizer, insecticides, and machinery) in the agricultural sector was an important element in the successful development of modern economies including such countries as Germany, Japan and the United States.

The classical model of Lewis-Fei-Ranis has also come under attack due to their assumption of zero marginal productivity of labour as a required condition for surplus of labour. This has been challenged by writers such as Viner (1957), Schultz (1964), Jorgenson (1961, 1967) and Sen (1966). Viner (1957) claims that "as far as agriculture is concerned, I find it impossible to conceive of a farm of any kind on which, other factors of production being held constant in quantity and even in form as well, it would not be possible by known methods, to obtain some addition to the crop by using additional labour in more careful section and planting the seed, more intensive weeding, cultivation, thinning, and mulching, more painstaking harvesting, gleaning, and cleaning the crop." Sen (1966) argues that the assumption of zero marginal productivity of labour is neither a necessary nor a sufficient condition for the existence of surplus labour. His reason is that even if it were shown that the marginal productivity of labour in agriculture was not zero but positive, it will not follow that there is no surplus of labour in a sense that any withdrawal of such labour from agriculture does not automatically imply a reduction in production. This can happen in such a case as, if a family worker is withdrawn from agriculture, then the

remaining family members increase their working efforts/hours so as to compensate any potential loss of production due to such labour withdrawal; and thus keep the level of production the same. Jorgenson (1967), in his neoclassical model of a dual economy, even went further by assuming that the marginal productivity of labour is always positive and that disguised unemployment is non-existent.

Another source of controversy is the assumption of a fixed real wage-rate in both agriculture and industry – from the point of view that industrial labour is available in unlimited amounts at a fixed real wage-rate – while increases in productivity are channeled into capital formation. This view is challenged also by the neoclassical school of thought. According to Jorgenson's model, the wage-rate is assumed to be variable rather than fixed. The industrial real wage rate rises over time, depending on the rate of technical progress in both sectors (which affects the internal terms of trade) and the rate of capital accumulation. Furthermore, wage rates in the agricultural sector are assumed to be proportional to those in the advanced sector. It follows that the wage differential is assumed to be proportional to the manufacturing wage-rate and is stable in the long run (Jorgenson, 1961). Guha (1969) questions the plausibility of the assumption of a constant real wage rate in the Fei-Ranis model and shows that the wage rate cannot be constant unless the marginal propensity to consume food is zero, i.e. non-existence of Engel's Law.

Finally, in the classical model population growth is ignored or shunted aside as a qualification to the main argument. According to Lewis (1954), "population increase is not relevant either to the classical analysis, or to the analysis which follows [in this surplus model], unless it can be shown that the increase of population is caused by economic development and would not otherwise be so large". In the Fei and Ranis model (1961), population growth is treated as a known phenomenon exogenous to the model. On the other hand, Jorgenson (1961) treats population growth as endogenous and assumes that population growth is dependent on the supply of food per capita and the force of mortality, where the latter is assumed to be exogenously given. Sato and Niho (1971), assuming population growth as a function of per capita income, show that as long as the actual growth rate of the population exceeds the technologically determined maximum rate, capital accumulation and technical progress in industry will have a negligible role in assisting the economy

to escape stagnation. A successful development program thus must rely on an appropriate population policy and a sufficiently high rate of technical progress in agriculture.

6.3 Sample Countries and Data

The results presented in this paper are based on a sample of 99 developed and developing countries. The sample countries exclude city-states such as Singapore and Hong Kong, where agriculture has never played any significant role due to their small size and little natural endowment, as well as other small countries, i.e. those whose land area is less than 10,000 square kilometers. Also excluded are former and current socialist countries. Table 6.1 gives the list of the sample countries.

The data used in this paper are for the year 1995, except where noted otherwise. We use the log of relative labour productivity in agriculture versus other sectors in 1995 (LRLP95) as the dependent variable. Relative labour productivity (RLP) for each country in a particular year can be calculated using a simple formula:

$$RLP = [\frac{1-s}{s}]\,[\frac{a}{1-a}]$$

where s and a are the agricultural shares of output and employment, respectively. This follows Bourguignon and Morrisson (1998). The data on agricultural shares of output and labour are obtained from the World Bank World Development Indicators CDROM (1999) and Food and Agricultural Organization Statistics (FAOSTAT), respectively.

For the independent variables, we use several variables of interest obtained from various sources. They include indicators of economic development, human and financial capital, geography and the political system.

One explanatory variable is real gross domestic product per capita measured in current prices in 1995 and expressed in natural logs (LGDP95); and it is used as a proxy for the level of economic development for each country in that particular year. The data for in-

come levels is obtained from Heston, Summers and Aten's Penn World Table version 6.1 (PWT 6.1).[77]

As for the measure of human capital, data on educational attainment were obtained from Barro and Lee's educational data set.[78] The log of the average years of schooling of the population aged 15 and over for 1995 (LSCH95) is used rather than either school enrollment rates or adult literacy rates which have been commonly used in cross-country growth studies. As Barro and Lee (1996) suggest, one weakness of these traditionally used variables is that they do not correspond well to the stock of human capital that influences current decision about fertility, health and so on. In addition to the current stock of human capital, the log of the average years of schooling in 1970 is used as an indicator of the initial level of human capital (LSCH70) and to account for a possible lagged effect that human capital might have on the degree of dualism.

We use two types of financial measures to account for the level of financial development and capital accumulation. First, this paper uses the same set of indicators of the level of financial development to measure the functioning of the financial system as those used in Levine, Loayza and Beck (2000a, 2000b), though in this paper these variables are expressed in log form. The first measure is the log of liquid liabilities (in 1995), LLLY95, which measures the financial depth or the size of financial intermediaries in the economy. It equals liquid liabilities of the financial system (currency plus demand and interest-bearing liabilities of banks and nonbank financial intermediaries) divided by GDP.

The second measure of financial development is the log of the commercial-central bank credit ratio, which measures the degree to which the central bank versus commercial banks is allocating credit. This variable is denoted as LBANK95 and equals the ratio of bank credit divided by bank credit plus central bank domestic assets. It can be used to indicate the relationship between the types of financial intermediaries that are conducting financial intermediation and the dependent variable, relative labour productivity. The final variable is the log of private credit, LPRIV95, which indicates private credit by deposit

---

[77] Detailed description of the programs and data flow in the construction of PWT6.1 can be found in the Technical Notes at the Center for International Comparisons of the University of Pennsylvania (CICUP) website. Also discussion of the overall data can be found in the Summers and Heston (1991) paper.

[78] The data are drawn from the Barro and Lee (2001) updated version of the dataset which contains the data from 1960 to 2000. Further discussion of the data can also be found in Barro and Lee (1993 and 1996).

money banks and other financial institutions as a ratio to GDP. It is calculated as private credit extended by deposit money banks and other financial institutions divided by gross domestic product.

Another financial variable used in this paper is the log of gross domestic saving as a percentage of gross domestic product in 1995 (LGDS95) as well as its initial level in 1970 (LGDS70). The saving data are obtained from the World Bank's World Development Indicators CDROM (1999) and, according to the World Bank's definition, are calculated as the difference between gross domestic product and total consumption in a particular year. Gross domestic saving is used as a proxy for the level of capital accumulation in the economy. The classical model such as that of Lewis views saving as a vital element in the growth process because it determines the rate of investment and capital accumulation in the modern sector, which works to absorb the surplus of labour in the backward sector as well as to promote economic growth.

To determine the role of geography, we use data on geographical variables obtained from the Harvard University's Center for International Development dataset. The first set of geographical variables to investigate are various measures of coastal access because, as discussed in Malik and Temple (2004), proximity to a coast or river may be associated with an increased ability to develop manufacturing exports and export diversification. This would lead to a reduction in these countries' dependency on primary exports production and the share of agriculture in the economy. The geography variables to be examined include the log of the mean distance from the nearest coastal line (LDISTCR), the 1994 share of population within 100km of a coast (POP100KM), the 1994 share of population within 100km of a coast or navigable river (POP100CR), the proportion of a country's total area within 100km of the ocean or ocean navigable river (LND100CR), and a dummy variable for landlocked countries (LANDLOCK). Other geographical factors we examine include a dummy for tropical countries (TROPICAL) – defined as those countries where the absolute value of latitude is less than or equal to 23, the latitude distance from equator expressed in natural log (EQDIST), and a dummy for African countries (AFRICA).

In examining the role of the political system, we use an index of political freedom as an indicator of the level of political openness of the country. The type of political regime for each individual country can be classified using the indexes for political rights and civil

liberties estimated by Freedom House. These indexes range from 1 to 7, with 1 indicating the most rights and liberties 7 the least. As defined by Freedom House, "Political rights enable people to participate freely in the political process, [and] civil liberties are freedom to develop views, institutions, and personal autonomy apart from state" (Ryan, 1994). The value for an index of political freedom used in our analysis for a country in a particular year is calculated as an average of the values of political rights and civil liberties which are obtained from the Freedom House data base. In addition, the index for political freedom used in our analysis is taken as the average value between 1994 and 1996 and expressed in natural logs (LPF9496).

Other explanatory variables used in our study include the log of agricultural machinery, measured as the number of tractors per hectare of arable land and averaged over the period 1980-95 (LMACH) and the log of the average fertilizer consumption between 1980-95 (100 grams per hectare of arable land) (LFZER) – used as proxies for agricultural capital inputs. We also use log of the population growth rate in 1995 (LPOPGR), log of the age-dependency ratio (LAGERAT), log of a measure of a country's openness measured in current prices (LOPENC), log of foreign direct investment as a percentage of GDP (LFDIGDP) and a dummy for low income countries (LOWINC). The data for all of these variables, except for LOPENC, are obtained from the World Bank World Development Indicators while the data for LOPENC are from Heston et al.'s PWT 6.1.

For those variables with missing values, we fill in the missing values by using a regression-based imputation method so as to maximize the number of observations for our econometric analysis of the dataset. The idea is to generate the missing value for a variable based on the values of other variables which are present using a linear regression. However, we realize the limitation of this method in providing estimation of the missing values and the potential impact that errors may have on the actual econometric analysis, especially for those variables with a fairly large number of missing values. Thus, we only impute those variables with less than ten missing values in order to minimize the impact that these estimated values may have on our analysis. Those variables of interest with more than ten missing values are dropped. Table 6.2 gives variable description and data sources, while Table 6.3 shows simple correlations among the explanatory variables.

6.4 Empirical Results

In this section we will present the findings of our empirical exercise. Our dependent variable is the log of relative labour productivity (LRLP95) measured in 1995. The analysis is conducted using data from 99 developing and developed countries and with a set of 24 variables. The Bayesian Model Averaging (BMA) method is used to evaluate sets of possible candidate explanatory variables. The Ordinary Least Squares (OLS) method is also applied to evaluate the top ten models suggested by the BMA exercise.

6.4.1 BMA Results

We use the model averaging method to determine the variables to be included in the models by calculating their respective posterior probabilities of inclusion. We also provide a sign certainty index, which is evaluated based on the sum of posterior model probabilities for all models in which a variable acts in a given direction (e.g. negative). The sign certainty index enables us to see the likely direction of the relationship between the dependent variable and various explanatory variables under consideration. The results of the Bayesian Averaging exercises are shown in Table 6.4.

The first quantity to consider is the level of economic development as indicated by income per capita. A common characteristic of poor countries is the coexistence of a large agricultural sector with a small and active industrial sector. There is generally a lack of integration between these sectors as reflected in a large intersectoral wage and productivity gap. Some studies on structural transformation of the economy, in particular that of Chenery and Syrquin (1975), show that as incomes rise this gap initially rises but falls later on once income reaches a certain level, namely there is an inverted-U relationship between income and the degree of economic dualism. A brief look at our data seems to support this view. Figure 6.1 displays a time series plot of the median values across countries of the relative labor productivity between 1961 and 1995. This plot shows an initially rising and then declining trend of the extent of dualism as the economy develops over time.

To empirically test this hypothesis on our sample data from 99 countries, we use the log of per capita GDP of each country in 1995 (measured in current year international prices) – LGDP95 – as a measure of a country's level of development and include it as

an explanatory variable in the BMA in which the log of relative labour productivity in 1995 is taken as the dependent variable. GDP per capita is a good measure of the level of development in the sense that it is correlated with most of the processes occurring with development, and thus it may capture the net effect of these processes as observed in cross country experience. To be consistent with other studies on structural transformation, and allow for a non-linear relationship, we also include the squared value of log income per capita ($LGDP^2$) as an explanatory variable.

The results of our empirical examination indicate a possible relationship between income level and the degree of dualism. In columns (1)-(3) and (5) the level of economic development as proxied by log of GDP per capita and its squared value are found to have a fairly and consistently high posterior probability of being included in the models. More importantly, the coefficients of the per capita income and of its squared value have the appropriate opposite signs as reflected in their sign certainty indexes – positive and negative, respectively – and thus support the hypothesis of an inverted U-shaped relationship between level of economic development and degree of dualism.

As explained above, to determine the possible explanatory role of the level of financial development in a country we include three indicators of financial development – LLLY95, LBANK95 and LPRIV95 – as explanatory variables in our exercises.

Columns (1)-(5) of Table 6.4 present a rather ambiguous picture of the relationship between the level of financial development and the degree of dualism. Out of the three financial development variables, only LBANK95 receives a high posterior probability of being included, while the other two variables do not. Furthermore, the fact that the sign certainty index for LBANK95 is positive does present a dilemma in the interpretation of the relationship because as an index of the level of financial development we would expect the relationship to be negative. However, one possible explanation for this positive relationship can be found if we consider LBANK95 as a measure of the importance of private banks in allocating resources in the economy. In such a context, private banks can contribute to dualism through their resource allocation decisions, since they are more likely to favour the modern sector over the traditional sector because of the relatively low risk as well as potentially higher returns associated with lending to the former.

This study also uses gross domestic saving in 1995 (LGDS95) and 1970 (LGDS70) to proxy for the current and initial levels of capital accumulation in the economy, respectively. As discussed above, saving is a very important component of economic development due to its nature as a channel for mobilizing resources necessary for investment and growth. The results of the BMA exercises show that the relationship between the current level of domestic savings and the degree of dualism is rather ambiguous. In column (3) we condition on the initial level of domestic saving (LGDS70) and find that it obtains a high posterior probability of inclusion and thus has a high explanatory power. This confirms the possibility of lagged effects that capital accumulation might have on the degree of structural integration in the economy. In addition the sign certainty index does point to a negative relationship between the two suggesting that capital accumulation does in the long run help improve the degree of structural integration in the economy.

To determine a possible role of human capital, we use average years of schooling for the total population aged 15 and over as a proxy for the stock of human capital –because education has a positive correlation with the productivity level of the labour force. A more educated labour force is generally assumed to be more productive than an illiterate one. We use both the current level (LSCH95) as well as the initial level (LSCH70) of human capital in our analysis to account for possible lagged effects. Similar to our experiment with gross domestic saving, we determine separately the role of the current and initial level of human capital. Columns (1), (3), (4) and (5) show that the current level of human capital does not seem to have any significant role in explaining the cross-country variation in the degree of dualism. In column (2) we condition on the initial level of human capital and, as in the case of gross domestic savings, we find it to have good explanatory power as shown by its fairly high posterior probability of inclusion. Again this lends support to the view that human capital improvement tends to have a long term effect on the structure of the economy rather than an immediate effect. However, the explanatory power of the initial level of human capital is significantly reduced when the initial level of saving (LGDS70) is included as shown in column (5).[79] Table 6.3 shows that the correlation between these two

[79] We also conduct exercises with various combinations of our explanatory variables and find that LSCH70 has consistently obtained a high posterior probability of inclusion when LGDS70 is absent, but the posterior probability is significantly reduced when the latter is present.

variables is fairly high, about 0.64. Thus a possible explanation for this phenomenon is the possible indirect association between the two variables via a third variable that might also be an influence on dualism.

Moving to the role of political openness, the results of our exercises support the proposition of a positive relationship between political freedom and dualism. Table 6.4 shows that the index of political freedom (LPF9496) obtains high posterior probabilities of inclusion in all our exercises. In column (4) when we exclude income, the explanatory power of the political freedom indicator has substantially increased, in which case it appears to pick up some of the effects of the level of income.

As for the role of geography, Bayesian results show that tropical and landlocked countries tend to have a highly dualistic structure. The posterior probabilities for tropical dummy are consistently high in all the experiments, while those for landlocked are also significant but to a lesser extent. Similarly, the dummy for African countries is also highly likely to be included. It obtains the highest posterior probability of inclusion in all of our experiments. In addition, the effects of these geographical factors are robust to the exclusion of income level in column (4). They are also robust to the addition of the initial level of human capital and saving, column (2) and (3) respectively.

Results also show that other variables appear to have only limited explanatory power as reflected by their low posterior inclusion probabilities. For instance, population growth rate (LPOPGR) has a consistently low posterior probability of inclusion in all the cases. This is also true for a number of other geographic, demographic and economic variables.

The Bayesian Model Averaging Method also yields posterior model probabilities which can be used to rank models in accordance with their explanatory power. In Table 6.5 we illustrate the structure of the top ten models with the highest posterior probabilities. The combined posterior probability of the top ten models is only about 36 percent, indicating that the extent of model uncertainty is considerable. This suggests that model selection through a more conventional method is likely to produce misleading results.

6.4.2 OLS Results

In order to investigate the magnitude of the effects different variables have on the degree of dualism, we also present the results of our Ordinary Least Square estimation of the top ten models chosen through BMA. Results are presented in Table 6.6.

The results confirm the relationship between income levels and the extent of dualism. The coefficients of log income and its square are statistically significant at the 5 percent level in all the cases. In addition, we find that the signs of their coefficients are as expected. Using simple mathematics we can estimate the turning point for the income level, where we expect the extent of dualism to decline after it reaches a certain level of income. The estimation of the turning points as well as the number of countries that lie to the left of these turning points are also presented in the table.

Among the three indicators of the level of financial development only LBANK95 has a significant and positive relationship with the dependent variable. Interpreted separately, this could mean that the larger the role of private sector banks in allocating financial resources, the more dualistic the nature of the economy. It may be that private banks would be more willing to skew their resource allocation towards the modern sector which has relatively low risk and potentially higher returns compared to the traditional sector. This may apply especially to a situation when capital funds are limited and competition for such limited funds is intense.

In the regressions which contain LGDS70 we find that the initial level of capital accumulation does have a statistically significant and negative relationship with later relative labour productivity. This supports our view of the lagged effects that capital accumulation may have on the extent of dualism in a country.

The results from OLS estimation indicate a significant and positive relationship between political freedom and dualism. This supports the view that a more politically free country is likely to be associated with a higher degree of dualism for some of the reasons explained above.

Table 6.6 also shows a strong explanatory role of geographical location. We find that all three dummy variables for landlocked, tropical and African countries have statistically significant and positive relationships with the dependent variable. We may say that isolated

countries with less access to external markets do seem to have a more dualistic structure of the economy.

With regards human capital, as Table 6.4 indicates, the role of our indicator of human capital is not strong whenever the initial level of saving is included in the regressions. Therefore, in order to examine the role of the human capital we conduct BMA experiments without including LGDS70. The OLS results of some of the best models from these BMA experiments are presented in Table 6.7. We find that the initial level of human capital, LSCH70, does have a significantly negative relationship with the extent of dualism, which indicates a possible lagged effect that human capital exerts on the degree of structural integration of an economy. As for other explanatory variables, the results in Table 6.7 are consistent with those presented in Table 6.6 with regards to the significance and the direction of their relationship with the dependent variable.

6.5 Conclusion

In this paper we seek to explain some of the underlying factors contributing to cross-country differences in the extent of structural integration, that is, to understand why some countries are more dualistic than others. This is an issue that has not been extensively investigated before.

Bayesian Model Averaging methods are employed because of their known advantages over more conventional approaches, especially when dealing with the issues of model uncertainty in a situation like ours, where there are a large number of possible candidate explanatory variables and supporting theories are weak.

We find an inverted U-shape relationship between the extent of dualism and the income level, which is consistent with some previous work. Our results, however, show a rather ambiguous explanatory role for the level of financial development.

The analysis indicates important long run effects of human capital and physical capital accumulation in promoting growth and structural integration of the economy. With regards to geography variables, we find that landlocked and tropical countries tend to be associated with a more dualistic structure. This might be because of their lack of access to

shipping routes for international trading, so these countries tend to concentrate their activities in agriculture and exporting primary products rather than promoting a modern sector that usually involves producing high value added goods for export and trade. In addition, our results show that African countries tend to exhibit a higher degree of dualism.

Lastly, a more politically open society is also found to be associated with a higher extent of dualism which may be due to a higher degree of intersectoral wage distortion. One possible explanation is that urban workers in these countries usually wield more bargaining power through various channels such as labour unions and thus are more able to secure higher wages than if wages were to be set by the market.

Table 6.1: List of Sample Countries

DZA Algeria	GIN Guinea	RWA Rwanda
AGO Angola	GNB Guinea-Bissau	SEN Senegal
ARG Argentina	GUY Guyana	SLE Sierra Leone
AUS Australia	HTI Haiti	ZAF South Africa
AUT Austria	HND Honduras	ESP Spain
BGD Bangladesh	ISL Iceland	LKA Sri Lanka
BEL Belgium	IND India	SWE Sweden
BLZ Belize	IDN Indonesia	SYR Syria
BEN Benin	IRN Iran	TZA Tanzania
BOL Bolivia	IRL Ireland	THA Thailand
BWA Botswana	ITA Italy	TGO Togo
BRA Brazil	JAM Jamaica	TUN Tunisia
BFA Burkina Faso	JPN Japan	TUR Turkey
BDI Burundi	JOR Jordan	UGA Uganda
CMR Cameroon	KEN Kenya	GBR United Kingdom
CAN Canada	KOR Korea, Republic of	URY Uruguay
CAF Central African Rep.	LSO Lesotho	VEN Venezuela
TCD Chad	MDG Madagascar	ZMB Zambia
CHL Chile	MWI Malawi	ZWE Zimbabwe
COL Colombia	MYS Malaysia	
ZAR Congo, Dem. Rep.	MLI Mali	
COG Congo, Republic of	MRT Mauritania	
CRI Costa Rica	MEX Mexico	
CIV Cote d'Ivoire	MAR Morocco	
DNK Denmark	MOZ Mozambique	
DOM Dominican Republic	NAM Namibia	
ECU Ecuador	NPL Nepal	
EGY Egypt	NLD Netherlands	
SLV El Salvador	NZL New Zealand	
GNQ Equatorial Guinea	NIC Nicaragua	
ETH Ethiopia	NER Niger	
FJI Fiji	NGA Nigeria	
FIN Finland	NOR Norway	
FRA France	PAK Pakistan	
GAB Gabon	PAN Panama	
GMB Gambia, The	PNG Papua New Guinea	
GER Germany	PRY Paraguay	
GHA Ghana	PER Peru	
GRC Greece	PHL Philippines	
GTM Guatemala	PRT Portugal	

Notes: This table shows names and country codes for the 99 sample countries used in this study. They are ranked in alphabetical order.

Table 6.2: Variable Description and Data Sources

VARIABLES	DESCRIPTION	SOURCES
LRLP95	Log of relative labor productivities in agriculture versus other sectors	The data used to calculate RLP, namely the agricultural shares of output and labor are from the World Bank Development Indicators and FAOSTAT, respectively.
LGDP95	Log of real gross domestic per capita, measured in constant price in 1995	Summers, Heston and Aten's PWT version 6.1
$LGDP^2$	Square of log of real gross domestic product in 1995	Ibid.
LLLY95	Log of liquid liabilities – which equals to liquid liabilities of financial system (currency plus demand and interest-bearing liabilities of banks and nonblank financial intermediaries) divided by GDP, in 1995.	Beck, Levine and Loayza (2000) financial dataset.
LBANK95	Log of BANK -- the ratio of bank credit divided by bank credit plus central bank domestic assets, in 1995.	Ibid.
LPRIV95	Log of PRIVATE – the ratio of private credits by deposit money banks and other financial institutions to GDP, in 1995.	Ibid.
LSCH95	Log of the average years of schooling for the population age 15 and over, in 1995	Barro and Lee's Educational Attainment Dataset, updated version (in 2001)
LSCH70	Log of the initial average years of schooling for the population age 15 and over, in 1970.	Ibid.
AFRICA	Dummy for African countries	World Bank World Development Indicators CD ROM (1999)
LOWINC	Dummy for low income countries	Ibid.
LGDS95	Log of gross domestic savings in 1995.	Ibid.
LGDS70	Log of the initial level of gross domestic savings, measured in 1970.	Ibid.
LDISTCR	Log of the mean distance from the nearest coastal line.	Harvard University's Center for International Development geography dataset.
POP100KM	1994 share of population within 100km of a coast	Ibid.

VARIABLES	DESCRIPTION	SOURCES
POP100CR	1994 share of population within 100km of a coast or navigable river.	Harvard University's Center for International Development geography dataset.
LND100CR	The proportion of a country's total area within 100km of the ocean or ocean navigable river.	Ibid.
LANDLOCK	Dummy for landlocked countries.	Ibid.
TROPICAL	Dummy for tropical countries.	Ibid.
EQDIST	The latitude distance from equator expressed in natural log.	Ibid.
LPF9496	Log of the average index of political freedom between 1994 and 1996.	Data for the two indexed used to compute the political freedom index, i.e. the indexes for political rights and civil liberties, are from Freedom House's database.
LMACH	Log of the average number of tractors per hectare of arable land over the period 1980-95.	World Bank World Development Indicators CD ROM (1999)
LFZER	Log of the average fertilizer consumption between 1980-95 (100 grams per hectare of arable land).	Ibid.
LAGERAT	Log of the age dependency ratio in 1995.	Ibid.
LPOPGR	Log of the rate of population growth in 1995.	Ibid.
LOPENC	Log of the measure of a country's openness measured in current prices in 1995.	Summers, Heston and Aten's PWT version 6.1

Table 6.3: Correlations between Explanatory Variables

	LGDP95	LGDP2	LBANK95	LLLY95	LPRIV95	AFRICA	LANDLOCK	TROPICAL	LOWINC	LGDS95	LGDS70	LSCH95	LSCH70
LGDP95	1.00												
$LGDP^2$	0.99	1.00											
LBANK95	0.55	0.54	1.00										
LLLY95	0.66	0.65	0.48	1.00									
LPRIV95	0.77	0.77	0.58	0.86	1.00								
AFRICA	-0.73	-0.71	-0.46	-0.62	-0.64	1.00							
LANDLOCK	-0.38	-0.37	-0.17	-0.31	-0.35	0.40	1.00						
TROPICAL	-0.61	-0.62	-0.38	-0.50	-0.47	0.50	0.12	1.00					
LOWINC	-0.82	-0.81	-0.47	-0.53	-0.60	0.76	0.36	0.42	1.00				
LGDS95	0.56	0.54	0.33	0.40	0.41	-0.42	-0.32	-0.32	-0.56	1.00			
LGDS70	0.61	0.60	0.18	0.37	0.46	-0.42	-0.36	-0.30	-0.53	0.46	1.00		
LSCH95	0.85	0.84	0.40	0.58	0.67	-0.65	-0.33	-0.51	-0.74	0.58	0.64	1.00	
LSCH70	0.78	0.79	0.33	0.49	0.62	-0.53	-0.37	-0.40	-0.68	0.45	0.63	0.88	1.00
LPOPGR	-0.77	-0.79	-0.40	-0.59	-0.62	0.52	0.22	0.60	0.54	-0.36	-0.46	-0.59	-0.60
LAGERAT	-0.88	-0.89	-0.53	-0.68	-0.75	0.72	0.39	0.58	0.76	-0.51	-0.53	-0.77	-0.76
LPF9496	-0.75	-0.77	-0.43	-0.50	-0.60	0.48	0.14	0.49	0.51	-0.31	-0.39	-0.59	-0.59
LMACH	0.85	0.85	0.42	0.65	0.67	-0.69	-0.35	-0.61	-0.74	0.56	0.56	0.78	0.72
LFZER	0.74	0.74	0.54	0.68	0.70	-0.74	-0.45	-0.48	-0.62	0.44	0.43	0.65	0.57
LDISTCR	-0.46	-0.45	-0.27	-0.35	-0.43	0.52	0.52	0.16	0.44	-0.22	-0.26	-0.34	-0.42
EQDIST	0.75	0.76	0.43	0.58	0.63	-0.53	-0.25	-0.81	-0.52	0.34	0.39	0.58	0.53
LND100CR	0.48	0.47	0.25	0.36	0.42	-0.52	-0.50	-0.17	-0.44	0.21	0.27	0.34	0.44
POP100CR	0.59	0.58	0.31	0.43	0.47	-0.59	-0.62	-0.28	-0.54	0.28	0.37	0.45	0.50
LNFDIGDP	0.18	0.16	-0.03	0.14	0.25	-0.17	-0.13	0.16	-0.31	0.09	0.16	0.22	0.35
LNOPENC	0.01	0.01	-0.11	0.13	0.12	0.05	-0.05	0.15	-0.09	0.12	0.06	0.11	0.14

	LPOPGR	LNAGERAT	LNPF9496	LNMACH	LNFZER	LNDISTCR	EQDIST	LND100CR	POP100CR	LFDIGDP	LOPENC
LPOPGR	1.00										
LAGERAT	0.82	1.00									
LPF9496	0.71	0.70	1.00								
LMACH	-0.72	-0.79	-0.67	1.00							
LFZER	-0.60	-0.74	-0.57	0.82	1.00						
LDISTCR	0.42	0.51	0.40	-0.46	-0.57	1.00					
EQDIST	-0.69	-0.66	-0.65	0.72	0.59	-0.35	1.00				
LND100CR	-0.47	-0.56	-0.43	0.45	0.56	-0.95	0.37	1.00			
POP100CR	-0.50	-0.63	-0.46	0.51	0.60	-0.85	0.46	0.90	1.00		
LFDIGDP	-0.03	-0.19	-0.18	0.18	0.08	-0.19	0.04	0.18	0.18	1.00	
LOPENC	0.08	0.02	-0.05	0.13	0.07	-0.19	0.04	0.16	0.15	0.48	1.00

Table 6.4: BMA Results

Dependent Variable	LNRLP95	LNRLP95	LNRLP95	LNRLP95	LNRLP95
Sample Countries	99	99	99	99	99
Variables	(1)	(2)	(3)	(4)	(5)
AFRICA	0.945 (+)	0.968 (+)	0.975 (+)	0.935 (+)	0.980 (+)
LANDLOCK	0.918 (+)	0.779 (+)	0.676 (+)	0.914 (+)	0.665 (+)
LBANK95	0.807 (+)	0.765 (+)	0.537 (+)	0.848 (+)	0.528 (+)
LPF9496	0.689 (+)	0.550 (+)	0.406 (+)	0.917 (+)	0.394 (+)
TROPICAL	0.441 (+)	0.440 (+)	0.383 (+)	0.555 (+)	0.366 (+)
LGDP95	0.279 (+)	0.300 (+)	0.536 (+)		0.543 (+)
$LGDP^2$	0.262 (-)	0.295 (-)	0.491 (-)		0.499 (-)
LFDIGDP	0.115	0.064	0.146	0.088	0.115
LOPENC	0.104	0.068	0.113	0.053	0.101
EQDIST	0.068	0.056	0.003	0.093	0.030
LOWINC	0.054	0.028	0.019	0.064	0.009
LMACH	0.019	0.025	0.016	0.023	0.012
LPRIV95	0.018	0.019	0.005	0.047	0.004
LGDS95	0.017	0.000		0.002	0.000
LGDS70			0.865 (-)		0.846 (-)
LSCH95	0.017		0.000	0.004	0.000
LSCH70		0.429 (-)			0.100 (?)
LDISTCR	0.015	0.000	0.000	0.036	0.000
LAGERAT	0.000	0.006	0.000	0.000	0.000
LPOPGR	0.000	0.000	0.000	0.002	0.000
LFZER	0.000	0.000	0.000	0.000	0.000
LND100CR	0.000	0.000	0.000	0.015	0.000
POPCR	0.000	0.000	0.041	0.000	0.028
LLLY95	0.000	0.000	0.000	0.019	0.000

Notes: The numbers reported in the table are the posterior probabilities of inclusion for each variable. The signs in the parentheses are the sign certainty indexes, which indicate the direction of the relationship between each variable and the dependent variable. It is based on the sum of posterior model probabilities for all the models in which a variable acts in a given direction (e.g. negative). For those numbers without the attached signs, it means that the sign of the relationship is uncertain.

Table 6.5: Structures of the Top Ten Models and Their Posterior Probabilities

Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
LGDP95	✓	✓				✓	✓			✓
$LGDP^2$	✓	✓				✓	✓			✓
LBANK95		✓	✓	✓	✓			✓	✓	
LGDS70	✓	✓		✓	✓	✓	✓		✓	✓
LANDLOCK	✓	✓	✓	✓	✓	✓		✓		✓
TROPICAL			✓		✓					
AFRICA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
LPF9496			✓	✓	✓			✓	✓	
LFDIGDP										✓
LOPENC						✓				
PMP	0.056	0.042	0.039	0.034	0.033	0.033	0.032	0.032	0.031	0.029

Notes: PMP stands for the posterior model probability
The sample consists of 99 developed and developing countries.

Table 6.6: OLS Estimates of the Top Ten Models suggested by BMA

Regression Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Observation	99	99	99	99	99	99	99	99	99	
CONSTANT	-12.353	-11.550	0.413	1.184	1.072	-12.716	-10.515	0.480	1.404	-13.710
	(-3.136)	(-2.959)	(2.998)	(3.291)	(3.002)	(-3.265)	(-2.658)	(3.495)	(4.007)	(-3.453)
LGDP95	3.389***	3.269***				3.701***	3.000***			3.699***
	(3.627)	(3.543)				(3.945)	(3.185)			(3.934)
$LGDP^2$	-0.201***	-0.198***				-0.219***	-0.178***			-0.219***
	(-3.664)	(-3.673)				(-3.979)	(-3.229)			(-3.967)
LBANK95		0.297*	0.440***	0.376**	0.414***			0.401**	0.367**	
		(1.952)	(2.965)	(2.536)	(2.826)			(2.665)	(2.434)	
LGDS70	-0.321***	-0.274**		-0.217**	-0.202**	-0.311***	-0.366***		-0.275***	-0.310***
	(-2.853)	(-2.420)		(-2.109)	(-1.993)	(-2.802)	(-3.235)		(-2.714)	(-2.790)
LANDLOCK	0.442**	0.429**	0.532***	0.393**	0.433**	0.429**		0.496***		0.436**
	(2.367)	(2.334)	(2.957)	(2.107)	(2.356)	(2.327)		(2.713)		(2.362)
TROPICAL			0.356**		0.334**					
			(2.223)		(2.112)					
AFRICA	0.890***	0.924***	0.568***	0.629***	0.513***	0.954***	0.947***	0.697***	0.734***	0.899***
	(4.315)	(4.531)	(3.191)	(3.669)	(2.894)	(4.617)	(4.513)	(4.061)	(4.381)	(4.410)
LPF9496			0.391***	0.390***	0.322**			0.468***	0.356***	
			(3.030)	(3.009)	(2.454)			(3.705)	(2.721)	
LFDIGDP										-0.067*
										(-1.773)
LOPENC						-0.241*				
						(-1.838)				
Adjusted-R^2	0.440	0.460	0.430	0.430	0.450	0.450	0.410	0.410	0.410	0.450
Hettest	0.14	0.28	0.08	0.05	0.25	0.25	0.08	0.005	0.11	0.13
Ovtest	0.51	0.20	0.05	0.004	0.08	0.48	0.59	0.008	0.002	0.54
β (LBANK95)	--	0.180	0.266	0.227	0.251	--	--	0.242	0.222	--
β (LGDS70)	-0.279	-0.239	--	-0.190	-0.176	-0.271	-0.319	--	-0.240	-0.270
β (LPF9496)	--	--	0.284	0.284	0.234	--	--	0.341	0.259	--
Turning Point	4584	3847				4674	4569			4653
Countries < TP	59	54				59	59			59

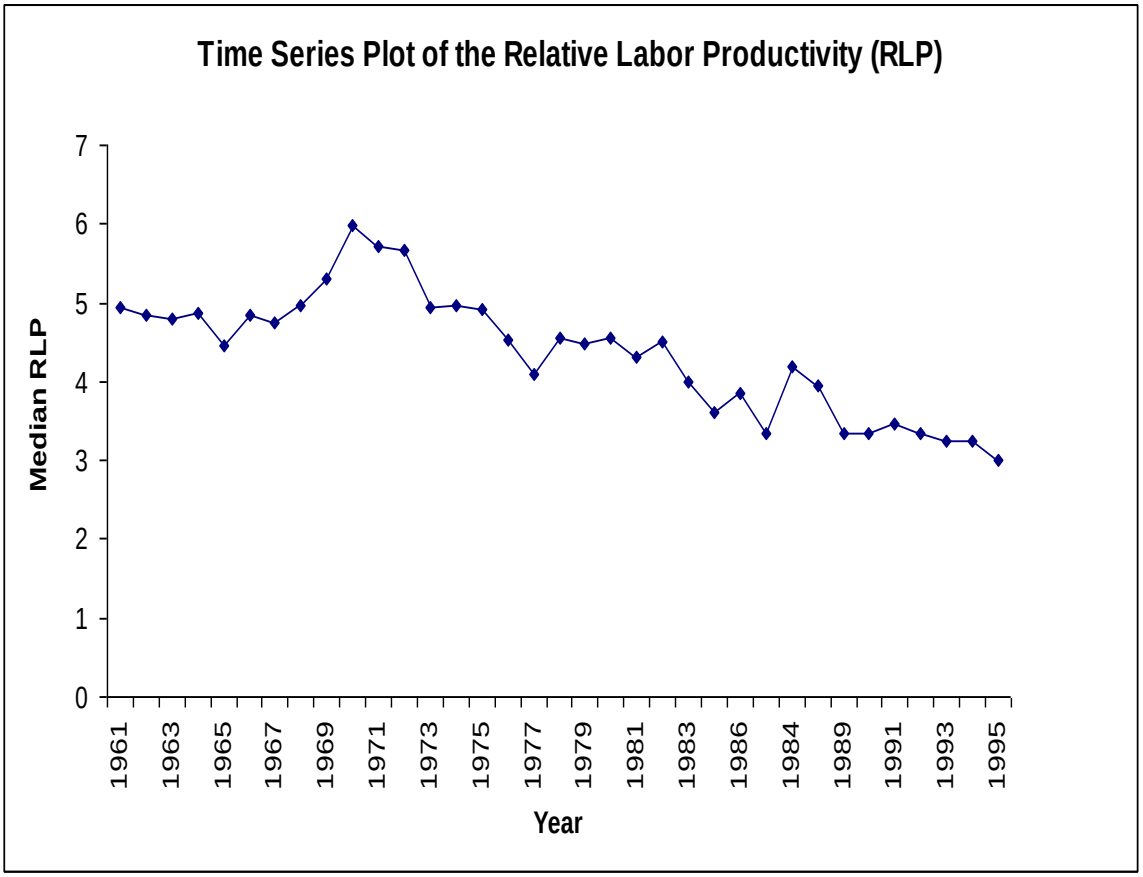
Notes: This table presents the OLS regression results with log of relative labor productivity in 1995 as the dependent variable (LRLP95). Numbers in the parentheses are t-statistics Hettest and Ovtest values represent the p-values for heteroscedasticity test and the Ramsey's RESET test for misspecification, respectively. β (*) is the beta value for the explanatory variable in the parenthesis. Turning Point is the dollar values of income where relative labor productivity begins to reverse course from rising to declining. Countries < TP is the number of countries that lie to the left of the turning point. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

T**able 6.7:** OLS Estimates of Some of the Best Models from BMA Experiments without LGDS70

Cross-country regression, log of relative labor productivity in 1995 as the dependent variable (LRLP95)					
Regression Model	(1)	(2)	(3)	(4)	(5)
Observation	99	99	99	99	99
CONSTANT	-9.324	1.076	0.713	0.422	-7.740
	(-2.271)	(5.924)	(2.851)	(3.068)	(-1.777)
LGDP95	2.491**				1.868*
	(2.513)				(1.687)
$LGDP^2$	-0.146**				-0.101
	(-2.450)				(-1.463)
LBANK95	0.335**	0.372**	0.437***	0.413***	0.332**
	(2.200)	(2.536)	(2.966)	(2.769)	(2.184)
LSCH70	-0.272**	-0.258**	-0.156		-0.226
	(-2.029)	(-2.611)	(-1.432)		(-1.606)
LANDLOCK	0.474**	0.400**	0.457**	0.518***	0.535***
	(2.552)	(2.131)	(2.452)	(2.885)	(2.872)
TROPICAL	0.269	0.413**	0.342**	0.431**	0.379**
	(1.474)	(2.608)	(2.144)	(2.548)	(2.001)
AFRICA	0.826***	0.558***	0.519***	0.517***	0.765***
	(3.853)	(3.060)	(2.879)	(2.852)	(3.580)
LPF9496			0.297**	0.343**	0.285
			(2.068)	(2.579)	(1.511)
LFDIGDP				-0.054	-0.045
				(-1.328)	(-1.051)
Adjusted-R^2	0.45	0.42	0.44	0.44	0.46
Hettest	0.10	0.11	0.04	0.09	0.10
Ovtest	0.40	0.06	0.09	0.15	0.43
β (LBANK95)	0.203	0.225	0.265	0.250	0.201
β (LSCH70)	-0.260	-0.247	-0.149	--	-0.217
β (LPF9496)	--	--	0.216	0.250	0.207

Notes: This table presents the Ordinary Least Squares estimates of the top five models suggested by Bayesian Model Averaging exercises. It differs from table 6.6 in that log of initial gross domestic savings in 1970 (LGDS70) is not included in the explanatory variable set. Numbers in the parentheses are t-statistics. Hettest and Ovtest values represent the p-values for heteroscedasticity test and the Ramsey's RESET test for misspecification, respectively. β (*) is the beta value for the explanatory variable in the parenthesis. *, **, *** indicate significance levels of 10, 5 and 1 percent respectively.

Figure 6.1: Time Series Plot of the Relative Labor Productivity



Notes: This figure displays a time series plot of the relative labor productivity (RLP). The data used to construct this figure are from a sample of 38 countries whose annual values for the relative labor productivity are available between 1961 and 1995. The values for RLP used in this case are the median for each year of these sample countries.

Chapter 7
Conclusion

This dissertation seeks to understand the sources of cross-country differences in the firm size distribution and the degree of structural integration in the economy. To achieve this objective, different empirical methods are used including Bayesian Model Averaging, GMM panel data estimation and Instrumental Variable estimation methods.

We employ different measures of the firm size distribution and of structural integration as the dependent variables. The SME share in the total official labour force (based on the official country definition of SME) is used as an indicator of the relative importance of this sector in the economy. In addition, we construct a new measure of the relative importance of the small enterprise sector based on available industrial data from the United Nations Industrial Development Organization (UNIDO). An advantage of this new variable is that it is measured in terms of the average number of employees per industrial establishment and thus enables us to avoid issues that may be associated with an arbitrary definition of different enterprise sectors.

With regards to the degree of structural integration (or structural dualism), the relative labour productivity in the agricultural sector versus other sectors is used as an indicator of the extent of dualism within an economy. It is calculated based on available annual data on agricultural shares of output and employment.

The studies in this dissertation yield several interesting findings. First, our results suggest that the relative importance of small and medium enterprises in the manufacturing sector and in the economy experience a decline in the long run as income levels rise. However, the relationship between the share of the SME sector and income is non-linear with the SME share initially rising, but then beginning to decline as income rises.[80] Such long-term declines in the relative importance of small and medium enterprises to the economy is in line with the findings of a number of previous studies including that of Mulhern

[80] On the other hand, when the UNIDO-based measure of the relative share of the small enterprise sector versus the large enterprise sector is used as the dependent variable, we do not find any significant relationship with the level of income. As explained earlier, a potential source of such an inconsistency is the differences between the two dependent variables, namely SMEOFF and SEMSELE, in terms of their construction and what they measure.

and Steward (2003) who attribute such a decline to the failures of small firms to improve their efficiencies and innovation, which limits their growth potential and, for some firms, even their ability to survive.

Second, our findings indicate less important roles of small enterprises in countries with higher levels of human capital. This, as Kremer (1993) suggests, may be because availability of higher skilled workers makes it easier for firms to specialize in more complex goods and acquire technologies that demand larger and more complicated production processes; and, hence, larger firm size.

Third, we also find that high quality and effective institutions are positively associated with the relative importance of small and medium enterprises in the economy. This finding shows that higher quality institutions may be associated with a business environment that is more friendly to small and medium sized firms and thus more conducive to their growth. In other words, politically stable countries with effective and efficient government institutions, low levels of corruption and higher degree of respect for the rule of law could be good breeding grounds for enterprises of small and medium sizes. Therefore, this finding may have valuable implications for policy makers who view SMEs as an important engine of growth and job creation.

However, it is important to recognize that these findings show only a potential correlation between institutional quality and the relative importance of the small and medium enterprise sector. In fact, there is insufficient evidence to suggest that institutions exert a causal effect on the small and medium enterprise sector.

We also find that not all regulations have positive effects on small and medium sized firms. Some regulations which impose additional costs on firms are found to have adverse effects, especially on smaller firms. For instance, regulations such as employment protection laws and employee's rights laws work to reduce the flexibility and ability of firms to take appropriate actions to adjust to market conditions, e.g. in hiring and firing workers. Such laws may have disproportionately large effects on small firms compared to larger firms because the latter has better abilities to absorb the costs resulting from such protective regulations.

Fourth, our analysis present a rather mixed picture of the relationship between different financial factors and the firm size distribution. On the one hand, our results show that

the level of financial development does not have any strong explanatory power for cross-country differences in the relative importance of the small and medium enterprise sector in the economy. This is not to say that financial development, which improves the functionality and efficiency of the financial system in allocating credit as well as increasing the availability of credit in the market, has no benefit for small and medium enterprises. Instead, the results may suggest that the development of the financial system may provide similar benefits to enterprises of all sizes, rather than favoring any one particular sector over the others.

On the other hand, we find a disproportionate effect of capital account liberalization on small firms versus large firms, suggesting that smaller firms benefit more from the liberalization. This may be because generally large firms are the main benefactors under a repressed financial system through such policy arrangements as credit controls and credit rationing. Under these arrangements, most credit usually goes to larger and more politically connected firms at the expense of smaller firms. For these reasons, removal of such financial repression policies would eliminate such distortions in credit allocation and, thus, should benefit smaller firms more than their larger counterparts.

Fifth, the role of international trade in explaining the firm size distribution across countries is rather ambiguous. On the one hand, we find a negative and significant relationship between the export share in GDP, our proxy for a country's exposure to international trade, and Ayyagari et al. measure of the share of the small and medium enterprise sector in the economy. This finding implies that average firm size tends to be large in countries with more exposure to international trading activities. On the other hand, the relationship between a country's exposure to external trade has a significantly positive relationship with the UNIDO-based relative size of the small enterprise sector versus the large enterprise sector, suggesting that small firms tend to have bigger roles than large firms under an open trade regime.

A possible explanation for this inconsistency rests on the differences between the two measures of small and medium enterprises – SMEOFF and SEMSELE – in terms of their construction. Another potential reason for the inconsistency in the findings concerns the extent of interfirm cooperation within a country. In countries such as Korea, there exists dynamic interfirm cooperation between firms of different sizes through such arrangements

as sub-contracting networks. Under such an arrangement, large firms rely heavily on small firms for supplies of intermediate products. Therefore, in these countries, expansion in international trading activities is likely to increase the role of small firms instead of reducing it. This is consistent with the finding when the relative size of the small enterprise sector, SEMSELE, is used as the dependent variable. Nevertheless, due to insufficient information on the extent of interfirm cooperation in each individual country, we do not know how many countries in our sample there are in which such interfirm cooperation plays a significant role. Thus, it is impossible to assess to what extent our results may be affected by such a factor.

In addition, we find that a country's relative geographical isolation, as indicated by its air distance to the closest major international port, is associated with an economy dominated by small firms. This is as expected because a country's geographical isolation from the external market would imply higher costs for firms in that country in conducting any business activities with the outside world, and thus limit their abilities to exploit larger external markets and reduce their abilities to grow.

Finally, a country's degree of de-industrialization, as proxied by the share of the service sector in GDP, is found to have a positive relationship with the relative share of small firm sector versus the large firm sector in the economy. A plausible explanation for this is that recent changes in the nature of demand in some countries from standardized goods to more specialized products and services have induced the movement away from large-scale manufacturing towards smaller industries which are capable of meeting such new demand requirements. As a consequence, the share of the small firm sector increases. This phenomenon has been observed in a number of developed countries, in particular those countries in which the service industry has taken over from manufacturing as the major source of employment and value added in the economy.

It is, nevertheless, important to recognize potential limitations of the studies of the firm size distribution in this dissertation. A potential limitation is due to problems associated with a lack of a coherent measure of the relative importance of small and medium enterprises, as indicated by a low correlation between the two measures used in this study, namely the Ayyagari et al. measure of the small and medium enterprise sector share (SME-OFF) and the UNIDO-based measure of the relative size of the small enterprise sector

versus the large enterprise sector (SEMSELE). This problem has led to a number of inconsistencies in our findings. Although these two measures have several advantages over other measures of SMEs, for reasons discussed earlier in this dissertation, such inconsistent results may illustrate the sensitivity of empirical exercises to changes in the set of measures of the relative importance of small and medium enterprises.

Another possible limitation to this dissertation concerns the extent of data coverage. Specifically, the data used for this dissertation only cover the period up to 1999 for the Ayyagari et al. measure (SMEOFF) and to 1996 for the UNIDO-based measure (SEMSELE). As a result, more recent evolutions of the firm size distribution are not taken into consideration. In light of recent evidence from some developed countries that suggest a revival of the SME role in the economy, it would be interesting to reexamine our findings once more up-to-date data become available.

With regards to the determinants of the degree of structural integration, we find that the extent of structural dualism – i.e. a lack of integration – has an inverted U-shape relationship with the level of income. Specifically, during the development process, countries are likely to become more dualistic at an earlier stage of development but become less so at a later stage of development, as incomes reach certain levels.

The degree of dualism is also found to be negatively related to the initial level of human capital, which may suggest potential lagged effects of the improvements in the quality of human resource on the extent of intersectoral integration in the economy. Likewise, empirical results indicate that physical capital accumulation helps reduce structural dualism and thus improves structural integration in the economy in the long run. We also find that politically freer countries tend to be more dualistic, perhaps due to a higher possibility of wage distortion produced by the wage bargaining power of interest groups such as labour unions.

BIBLIOGRAPHY

Abiad, A. and Mody, A. (2003), *Financial Reform: What Shakes It? What Shapes It?,* IMF Working Paper, International Monetary Fund, Washington, D.C.

Acemoglu, D., Johnson, S. and Robinson, J. (2001) 'The Colonial Origins of Comparative Development: An Empirical Investigation', *American Economic Review,* 91(5), 1369-1401.

Acs, Z. and Audretsch, D. (1990), *Small Firms in the 1990s,* in The Economics of Small Firms: A European Challenge, Norwell, MA, Kluwer Academic.

Acs, Z. and Audretsch, D. (1990), *The Economics of Small Firms: A European Challenge,* Kluver Academic Publishers.

Aghion, P. and Bolton, P. (1997) 'A Theory of Trickle-Down Growth and Development', *The Review of Economic Studies,* 64(2), 151-172.

Ahluwalia, M. (1976) *Inequality, Poverty and Development,* World Bank, Washington, D.C.

Alchian, A. and Demsetz, H. (1972) 'Production, information Costs, and Economic Organization', *American Economic Review,* 62, 777-795.

Alexander, G., Eun, C. and Janakiramanan, S. (1987) 'Asset Pricing and Dual Listing on Foreign Capital Markets: A Note', *Journal of Finance,* 42(1), 151-158.

Anderson, D. (1982) 'Small Industry in Developing Countries: A Discussion of Issues', *World Development,* 10, 11.

Anderson, B. (1983), *Imagined Communities,* Verso, London.

Aoki, M. (1988), *Information, Incentives and Bargaining in the Japanese Economy,* Cambridge: Cambridge University Press.

Arellano, M. and Bond, S. (1991) 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations', *Review of Economic Studies,*58, 277-297.

Arellano, M. and Bover, O. (1995) 'Another Look at the Instrumental Variable Estimation of Error-Components Models', *Journal of Econometrics,* 68, 29-52.

Audretsch, D. (2000), *The Economic Role of Small- and Medium-Sized Enterprises: The United States,* Institute for Development Strategies, Indiana University, Indiana.

Audretsch, D., Klomp, L. and Thurik, R (2002), *Gibrat's Law: Are the Services Different?*, ERIM Report Series 2002-04.

Ayyagari, M., Beck, T. and Demirguc-Kunt, A. (2003), *Small and Medium Enterprises Across the Globe: A New Database,* World Bank, Washington, D.C.

Banerjee, A. and Newman, A. (1993) 'Occupational Choice and the Process of Development', *The Journal of Political Economy,* 101(2), 274-298.

Banjeri, R. (1978) 'Average Size of Plants in Manufacturing and Capital Intensity: A Cross-Country Analysis by Industry', *Journal of Development Economics,* 5, 155-166.

Banerji, A. and Ghanem, H. (1997) 'Does the Type of Political Regime Matter for Trade and Labor Market Policies?', *The World Bank Economic Review,* 11(1), 171-193.

Barro, R. (1991) 'Economic Growth in a Cross-Section of Countries', *The Quarterly Journal of Economics,* 106, 407-443.

Barro, R. and Lee, J. (1996) 'International Measures of Schooling Years and Schooling Quality', *The American Economic Review,* 86(2), 218-223.

Barro, R. and Lee, J. (2001) 'International Data on Education Attainment: Updates and Implication', paper can be found via the internet at *http://econ.korea.ac.kr/prof/jwlee/files/lee%20oep(final)1.dat.pdf*

Baum, C., Schaffer, M. and Stillman, S. (2003) 'Instrumental Variables and GMM: Estimating and Testing', *Stata Journal,* 3(1), 1-31.

Baumol, W. et al. (1982), *Contestable Markets and the Theory of Industry Structure*, New York: Harcourt Brace Jovanovich.

Beck, T., Levine, R. and Loayza, N. (2000a) 'Financial Intermediation and Growth: Causality and Causes', *Journal of Monetary Economics,* 46, 31-77.

Beck, T., Levine, R. and Loayza, N. (2000b) 'Finance and the Sources of Growth', *Journal of Financial Economics*, 58, 261-300.

Beck, T., Demirgüç-Kunt, A. and Maksimovic, V. (2002), *Financial and Legal Constraints to Firm Growth: Does Size Matter?,* World Bank Policy Research Paper No. 2784, World Bank, Washington, D.C.

Beck, T., Demirguc-Kunt, A. and Levine, R. (2003), *Small and Medium Enterprises, Growth, and Poverty: Cross-Country Evidence*, World Bank Policy Research Working Paper 3178, World Bank, Washington, D.C.

Beck, T., Demirguc-Kunt, A. and Maksimovic, V. (2003), *Financial and Legal Institutions and Firm Size,* World Bank Policy Research Working Paper 2997, World Bank, Washington, D.C.

Beck, T. and Dermirguc-Kunt, A. (2004), *SMEs, Growth and Poverty: Do Pro-SME Policies Work?,* World Bank Development Research Group, World Bank, Washington, D.C.

Beck, T., Demirguc-Kunt, A., Laeven, L. and Levine, R. (2004), *Finance, Firm Size, and Growth*, NBER Working Paper No. 10983, National Bureau of Economic Research, Cambridge.

Berry, A. and Mazumdar, D. (1991) 'Small-Scale Industry in the Asian-Pacific Region', *Asian-Pacific Economic Literature,* 5(2), 35-67.

Biggs, T. and Shah, M. (1998), *The Determinants of Enterprise Growth in the Sub-Saharan Africa: Evidence from the Regional Program on Enterprise Development*, RPED Discussion Paper, World Bank, Washington, D.C.

Biggs, T. (2002), *Is Small Beautiful and Worthy of Subsidy? Literature Review*, IFC mimeo.

Bijmolt, T. and Zwart, P. (1994) 'The Impact of Internal Factors on Export Success of Dutch Small and Medium-Sized Firms', *Journal of Small Business Management*, 32(2).

Birch, D. (1981) 'Who Creates Jobs?,' *The Public Interest,* 65, 3-14.

Blundell, R. and Bond, S. (1998) 'Initial Conditions and Moment Restrictions in Dynamic Panel Data Models', *Journal of Econometrics*, 87(1), 115-143.

Bond, S. (2002), *Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice',* CEMMAP Working Paper CWP09/02.

Bourguignon, F. and Morrisson, C. (1998) 'Inequality and Development: the Role of Dualism', *Journal of Development Economics*, 57(2), 233-257.

Bowles, S. and Gintis, H. (1990) 'Contested Exchange: New Microfoundations for Political Economy of Capitalism', *Politics and Society,* 18, 165-222.

Brown, C., Medoff, J. and Hamilton, J. (1990), *Employers: Large and Small*, Cambridge, MA: Harvard University Press.

Bruno, G. (2004), *Approximating the Bias of the LSDV Estimator for Dynamic Unbalanced Panel Data Models*, CESPRI Working Paper No. 159, CESPRI, Milan, Italy.

Bruno, G. (2005), *Estimation and Inference in Dynamic Unbalanced Panel Data Models with a Small Number of Individual,* CESPRI Working Paper No. 165, CESPRI, Milan, Italy.

Brusco, S. (1982) 'The Emilian Model: Productive Decentralization and Social Integration', *Cambridge Journal of Economics*, 6, 167-184.

Brynjolfsson, E. (1994) 'Information Assets, Technology and Organization,' *Management Science,* 40(12), 1645-1662.

Bun, M. and Kiviet, J. (2003) 'On the Diminishing Returns of Higher Order Terms in Asymptotic Expansion of Bias', *Economics Letters,* 79, 145-152.

Cabral, L. and Mata, J. (2003) 'On the Evolution of the Firm Size Distribution: Facts and Theory', *American Economic Review,* 93(4), 1075-1090.

Canning, D., Fay, M. and Perroti, R. (1994) 'Infrastructure and Growth', in M. Baldassarri, M. Paganaetto and E. Phelps (Eds.), *International Differences in Growth Rates,* New York: St. Martin's Press.

Canning, D. and Pedroni, P. (1999) 'Infrastructure and Long Run Economic Growth', a Paper Presented at the 1999 Economic Society Summer Meeting', Madison, Wisconsin, USA.

Carisson, B. (1989) 'Flexibility and the Theory of the Firm', *International Journal of Industrial Organization,* 7, 179-203.

Caves, R. and Uekusa, M. (1976), *Industrial Organization in Japan,* Brookings Institution, Washington, DC.

Caves, R. (1998) 'Industrial Organization: New Findings on the Turnover and Mobility of Firms', *Journal of Economic Literature,* 36, 1947-1983.

Chamberlain, G. (1984) 'Panel Data', in Z. Griliches and M. Intriligator (Eds.), *Handbook of Economics, Vol. II,* Amsterdam, North Holland.

Chandler, A. (1966), *Strategy and Structure,* New York: Doubleday and Company.

Chenery, H. and Syrquin, M. (1975), *Pattern of Development 1950-1970,* London: Oxford University Press.

Chenery, H. (1979), *Structural Change and Development Policy,* London: Oxford University Press.

Chinn, M. and Ito, H. (2002), *Capital Account Liberalization, Institutions and Financial*

Development: Cross Country Evidence, NBER Working Paper No. 8967, National Bureau of Economic Research, Cambridge.

Cho, Y. (1984), *On the Liberalization of the Financial System and Efficiency of Capital Accumulation under Uncertainty*, Stanford: Stanford University, Ph.D. Thesis.

Chowdhry, B. and Nanda, V. (1991) 'Multimarket Trading and Market Liquidity', *Review of Financial Studies*, 4(3), 483-511.

Clark, X., Dollar, D. and Micco, A. (2004) *Port Efficiency, Maritime Transport Costs and Bilateral Trade*, NBER Working Paper 10353, National Bureau of Economic Research, Cambridge.

Coase, R. (1937) 'The Nature of the Firm', *Economica*, 4, 386-405.

Davis, S., Haltiwanger, J. and Schuh, S. (1993) 'Small Business and Job Creation: Dissecting the Myth and Reassessing the Facts', *Business Economics*, 29, 13-21.

Davis, S. and Henrekson, M. (1997), *Explaining National Differences in the Size and Industry Distribution of Employment*, IUI Working Paper Series No 482.

Demirguc-Kunt, A. and Maksimovic, V. (1999) 'Institutions, Financial Markets and Firm Debt Maturity', *Journal of Financial Economics,* 54(3), 295-336.

Dixit, A. (1970) 'Growth Patterns in a Dual Economy', *Oxford Economic Papers*, 22(2), 229-234.

Djankov, S., La Porta, R., Lopez-De-Silanes, F. and Shleifer, A. (2002), ' The Regulation of Entry', *The Quarterly Journal of Economics,* 117(1), 1-37.

Doi, N. and Cowling, M. (1998) 'The Evolution of Firm Size and Employment Share Distribution in Japanese and U.K. Manufacturing: A Study of Small Business Presence,' *Small Business Economics,* 10, 238-292.

Domowitz, I., Glen, J. and Madhavan, A. (1998) 'International Cross-Listing and Order Flow Migration: Evidence from an Emerging Market', *Journal of Finance,* 53(6), 2001-2027.

Dunne, T., Roberts, M. and Samuelson, L. (1989a) 'The Growth and Failure of U.S. Manufacturing Plants', *Quarterly Journal of Economics,* 104(4), 671-698.

Dunne, T., Roberts, M. and Samuelson, L. (1989b) 'Plant Turnover and Gross Employment Flows in the U.S. Manufacturing Sector', *Journal of Labor Economics,* 7(1), 48-71.

Dunne, P. and Hughes, A. (1990), *Small Business: An Analysis of Recent Trends in Their*

Relative Importance Growth Performance in the UK with Some European Comparisons, Small Business Research Center, Working Paper No. 1, Department of Applied Economics, University of Cambridge.

Easterly, W. and Rebelo, S. (1993) 'Fiscal Policy and Economic Growth: An Empirical Investigation', *Journal of Monetary Economics,* 32(3), 417-458.

Edison, H., Klein, M., Ricci, L. and Slok, T. (2002), *Capital Account Liberalization and Economic Performance: A Review of the Literature,* mimeo, International Monetary Fund, Washington, D.C.

Evans, D. (1987) 'Tests of Alternative Theories of Firm Growth', *Journal of Political Economy*, 95, 657-674.

Evans, D. and Jovanovic, B. (1989) 'An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." *Journal of Political Economy*', 97(4), 808–827.

Fay, M. and Perotti, R. (1994) 'Infrastructure and Growth', in M. Baldassarri, M. Paganaetto and E. Phelps (Eds.), *International Differences in Growth Rates,* New York: St. Martin's Press.

Fazzari, S., Hubbard, R. and Petersen, B. (1988) 'Financing Constraints and Corporate Investment', *Brookings Papers on Economic Activity,* 1, 141–195.

Fernandez, C., Ley, E. and Steel, M. (2001) 'Benchmark Priors for Bayesian Model Averaging', *Journal of Econometrics*, 100, 381-427.

Fonseca, R. and Utrero, N. (2006), *Employment Protection Laws, Barriers to Entrepreneurship, Financial Markets and Firm Size,* RAND Labor and Population Program Working Paper 454, RAND Corporation, Cambridge.

Freeman, D. (1983) 'A Note on Screening Regression Equations', *The American Statistician*, 37(2), 152-155.

Fry, M. (1995), *Money, Interest, and Banking in Economic Development,* Maryland: The John Hopkins University Press.

Galor, O. and Zeira, J. (1993) 'Income Distribution and Macroeconomics', *Review of Economic Studies*, 60, 35-52.

Georgescu-Roegen, N. (1972) 'Process Analysis and the Neoclassical Theory of Production', *American Journal of Agricultural Economics*, 54, 279-294.

Geroski, P. (1995) 'What Do We Know About Entry?', *International Journal of Industrial Organization,* 13, 421-440.

Gertler, M. and Gilchrist, S. (1991), *Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms*, NBER Working Paper No. 3892, National Bureau of Economic Research, Cambridge, MA.

Gibrat, R. (1931), *Les Inegalite Economiques*, Libraire du Recueil Sirey.

Glaeser, E., La Porta, R., Lopez-de-Silanes, F. And Shleifer, A. (2004) 'Do Institutions Cause Growth?', *Journal of Economic Growth,* 9(3), 271-303.

Godfrey, L. (1999) 'Instrument Relevance in Multivariate Linear Models', *Review of Economics and Statistics,* 81(3), 550-552.

Goodman, E. and Bamford, J. (1989), *Small Firms and Industrial Districts in Italy*, New York, Routledge.

Granger, C., King, M. and White, H. (1995) 'Comments on Testing Economic Theories and the Use of Model Selection Criteria', *Journal of Econometrics,* 67, 173-187.

Guha, A. (1969) 'Accumulation, Innovation, and Growth under Conditions of Disguised Unemployment', *Oxford Economic Papers,* 21(3), 360-372.

Guiso, L., Sapienza, P. and Zingales, L. (2002), *Does Local Financial Development Matter?*, NBER Working Paper No. 8922, National Bureau of Economic Research, Cambridge.

Hahn, J. and Hausman, J. (2002) 'Notes on Bias in Estimators for Simultaneous Equation Models', *Economics Letters,* 75(2), 237-241.

Hall, B. (1987) 'The Relationship Between Firm Size and Firm Growth in the US Manufacturing Sector', *Journal of Industrial Economics,* 35(4), 583-606.

Hall, R. and Jones, C. (1999) 'Why Do Some Countries Produce Much More Output per Worker than Others?', *Quarterly Journal of Economics,* 114(1), 83-116.

Hargis, K. (2000) 'International Cross-Listing and Stock Market Development in Emerging Economies', *International Review of Economics and Finance,* 9(2), 101-122.

Henrekson, M. (1992), 'Sweden: Monetary and Financial System,' in P. Newman, M. Milgate and J. Eatwell (Eds.), *The New Palgrave Dictionary of Money and Finance,* London: MacMillan.

Henrekson, M. and Johansson, D. (1999) 'Institutional Effects on the Evolution of the Size Distribution of Firms,' *Small Business Economics,* 12(1), 11-23.

Herrera, A. and Lora, E. (2005), *Why So Small? Explaining the Size of Firms in Latin America,* Oxford: Blackwell Publishing Limited.

Heston, A. and Summers, R. (1991) 'The Penn World Table (Mark 5): An Expanded Set of International Comparison, 1950-1988', *The Quarterly Journal of Economics,* 106(2), 327-368.

Heston, A., Summers, R. and Aten, B. (2002), *Penn World Table Version 6.1*, Center for International Comparisons at the University of Pennsylvania (CICUP).

Hoeting, J., Madigan, D., Raftery, A. (1997) 'Bayesian Model Averaging for Linear Regression Models', *Journal of the American Statistical Association, 92, 179-191.*

Hoeting, J., Madigan, D., Raftery, A. and Volinsky, C. (1999) 'Bayesian Model Averaging: A Tutorial', *Statistical Science*, 14(4), 382-417.

Holtz-Eakin, D., Newey, W. and Rosen, H. (1988) 'Estimating Vector Autoregressions With Panel Data', *Econometrica,* 56(6), 1371-1395.

Hopenhayn, H. (1992) 'Entry, Exit, and Firm Dynamics in Long Run Equilibrium', *Econometrica*, 60(5), 1127-1150.

Hu, M.-W. (1999) 'The Determinants of SMEs' Market Share in 1991 – Taiwan Manufacturers,' *Small Business Economics,* 12, 1-9.

Hughes, A. (1992), *The Problems of Finance for Small Businesses*, Small Business Research Centre, Working Paper No. 15, Department of Applied Economics, University of Cambridge.

IFC, *World Business Environment Survey*, International Finance Corporation, Washington, D.C.

IMF, *Annual Report on Exchange Arrangements and Exchange Restrictions*, International Monetary Fund, Washington D.C.

Inter-American Development Bank (2001), *Economic and Social Progress in Latin America: Competitiveness: The Business of Growth,* Inter-American Development Bank, Washington D.C.

Jayaratne, J. and Strahan, P. (1998) 'Entry Restrictions, Industry Evolution, and Dynamic Efficiency: Evidence from Commercial Banking', *Journal of Law and Economics,* 41(1), 239-273.

Jorgenson, D.W. (1961) 'The Development of a Dual Economy', *The Economic Journal,* 71(282), 309-334.

Jorgenson, D. (1967) 'Surplus Agricultural Labour and the Development of a Dual Economy', *Oxford Economic Papers,* 19(3), 288-312.

Jovanovic, B. (1982) 'Selection and Evolution of Industry', *Econometrica*, 50(3), 649-670.

Judson, R. and Owen, A. (1999) 'Estimating Dynamic Panel Data Models: A Guide for Macroeconomists', *Economics Letters,* 65, 9-15.

Kass, R. and Raftery, A. (1995) 'Bayes Factors', *Journal of the American Statistical Association,* 90(430), 773-795.

Kauffman, D., Kraay, A. and Zoido-Lobaton, P. (1999), *Government Matter,* World Bank Policy Research Department Working Paper 2196, World Bank, Washington, D.C.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2005), *Governance Matters IV: Governance Indicators for 1996-2004,* World Bank, Washington, D.C.

Kihlstrom, R. and Laffont, J. (1979) 'A General Equilibrium Entrepreneurial Theory of the Firm Based on Risk Aversion', *Journal of Political Economy,* 87, 719-748.

King, R. and Levine, R. (1993) 'Finance and Growth: Schumpeter Might be Right', *The Quarterly Journal of Economics,* 108(3), 717-737.

Kinyanjui, M., McCormick, D. and Ongile, G. (1997) 'Growth and Barriers to Growth Among Nairobi's Small and Medium Sized Garment Producers', *World Development,* 25, 1095-1110.

Kitching, G. (1982), *Development and Underdevelopment in Historical Perspective,* London: Metheun.

Kiviet, J. (1995) 'On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models', *Journal of Econometrics,* 68(1), 53-78.

Kiviet, J. (1999) 'Expectation of Expansions for Estimators in a Dynamic Panel Data Model; Some Results for Weakly Exogenous Regressors', in C. Hsiao, K. Lahiri, L.-F. Lee and M. Pesaran (Eds.), *Analysis of Panel Data and Limited Dependent Variables,* Cambridge: Cambridge University Press.

Knack, S. and Keefer, P. (1995) 'Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures', *Economics and Politics,* 7(3), 207-227.

Knight, F. (1965), *Risk, Uncertainty and Profit,* New York: Harper & Row.

Kremer, M. (1993) 'The O-Ring Theory of Economic Development', *Quarterly Journal of Economics, 108. 551-75.*

Krugman, P. (1978), *Interest Rate Ceilings, Efficiency and Growth: A Theoretical Approach,* New Haven: Yale University, Department of Economics.

Kumar, K., Rajan, R. and Zingales, L. (1999), *What Determines Firm Size?*, NBER Working Paper No. 7208, National Bureau of Economic Research, Cambridge.

Kuznets, S. (1971), *Economic Growth of Nations: Total Output and Production Structure,* Cambridge, MA: Harvard University Press.

La Porta, R., Lopez-de-Silanes, F.,Shleifer, A. and Vishny, R. (1997) 'Legal Determinants of External Finance', *Journal of Finance, 52(3), 1131-1150.*

La Porta, R., Lopez-de-Silanes, F.,Shleifer, A. and Vishny, R. (1998) 'Law and Finance', *Journal of Political Economy, 106(6), 1113-1155.*

La Porta, R., Lopez-de-Silanes, F.,Shleifer, A. and Vishny, R. (1999) 'The Quality of Government', *Journal of Law, Economics and Organization, 15(1), 222-279.*

Laeven, L. (2000), *Does Financial Liberalization Relax Financing Constraints on Firms?,* World Bank Policy Research Working Paper No. 2467, World Bank, Washington, D.C.

Lazonick, W. (1990), *Advantage on the Shop Floor*, Cambridge, MA: Harvard University Press.

Leamer, E. (1978), *Specification Searches. Ad-Hoc Inference with Non-Experiment Data,* New York, John Wiley.

Leamer, E. (1983) 'Let's Take the Con Out of Econometrics', *American Economic Review,* 73, 31-43.

Levine, R. and Renelt, D. (1992) 'A Sensitivity Analysis of Cross-Country Growth Regressions', *American Economic Review,* 82(4), 942-963.

Levine, R. (1997) 'Financial Development and Economic Growth: Views and Agenda', *Journal of Economic Literature,* 35(2), 688-726.

Levine, R. and Zervos, S. (1998) 'Stock Markets, Banks, and Economic Growth', *American Economic Review,* 88(3), 537-587.

Lewis, A. (1954) 'Development with Unlimited Supplies of Labor', *Manchester School,* 22, 139-192.

Liedholm, C. and Mead, D. (1998) 'The Dynamics of Micro and Small Enterprises in Developing Countries', *World Development,* 26, 61-74.

Little, I. (1987) 'Small Manufacturing Enterprises in Developing Countries', *World Bank Economic Review,* 1(2).

Little, I., Mazumdar, D. and Page, J. (1987), *Small Manufacturing Enterprises: A Comparative Study of India and Other Economies,* New York: Oxford University Press.

Loveman, G., Piore, M. and Segenberger, W., (1990), *The Re-Emergence of Small Enterprises: Industrial Restructuring in Industrialized Countries,'* International Labor Organization, Geneva.

Loveman, G. and Sengenberger, W. (1991), 'The Re-Emergence of Small-Scale Production: An International Comparison', *Small Business Economics,* 3(1), 1-37.

Lucas, R. (1978) 'On the Size Distribution of Business Firms', *Bell Journal of Economics,* 9, 508-523.

Madigan, D. and Raftery, A. (1994) 'Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window', *Journal of the American Statistical Association,* 89, 1535-1546.

Madigan, D. and York, J. (1995) 'Bayesian Graphical Models for Discrete Data', *International Statistical Review,*63, 215-232.

Malik, A. and Temple, J. (2004), *Geography, Institution and Output Volatility,* Working Paper, Department of Economics, University of Bristol.

Mansfield, E. (1962) 'Entry, Gibrat's Law, Innovation and the Growth of Firms', *American Economic Review,* 52(5), 1023-1051.

Marshall, A. (1948), *Principles of Economics,* 8th Edition, New York: Macmillan.

Marshall, M. and Jaggers, K. (2002), *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002,* project information can be found via the internet at *www.bsos.umd.edu/cidcm/polity.*

Marx, K. (1976), *Capital,* Volume 1, Harmondsworth: Penguin.

Mauro, P. (1995) 'Corruption and Growth', *Quarterly Journal of Economics,* 110(3), 681-712.

Mayint, H. (1971), *Economic Theory and the Underdeveloped Countries,* London: Oxford University Press, Inc.

Mazumdar, D. (2001) 'Small-Medium Enterprise Development in Equitable Growth and Poverty Alleviation', a Paper Presented at the Asia and Pacific Forum on Poverty: Reforming Policies and Institutions for Poverty Reduction, Held at the Asian Development Bank, Manila.

McGuinness, T. (1994) 'Markets and Managerial Hierarchies', in G. Thompson et al. (Eds.), Markets, Hierarchies and Networks, London, UK: Sage.

McKinnon, R. (1973), *Money and Capital in Economic Development,* Washington, D.C.: Brookings Institution.

Meier, G. (1989), *Leading Issues in Economic Development,* New York: Oxford University Press.

Micco, A. and Serebrisky, T. (2004), *Infrastructure, Competition Regimes, and Air Transport Costs: Cross-Country Evidence,* Policy Research Working Paper Series 3355, The World Bank, Washington D.C.

Milgrom, P. and Roberts, J. (1990) 'The Economics of Modern Manufacturing: Technology, Strategy and Organization', *American Economic Review,* 80, 511-528.

Morawetz, D. (1974) 'Employment Implications of Industrialization in Developing Countries: A Survey', *The Economic Journal,* 84(335), 491-542.

Mulhern, A. and Steward, C. (2003) 'Long-Term Decline of Small and Medium Size Enterprise Share', *Small Business Economics,* 21, 215-228.

Myint, H. (1971), *Economic Theory and the Underdeveloped Countries,* London: Oxford University Press, Inc.

Nabli, M. and Nugent, J. (1992) 'Development of Financial Markets and the Size Distribution of Manufacturing Establishments: International Comparisons', *World Development,* 20(10), 1489-1499.

Nerlove, M. (1999) 'Properties of Alternative Estimators of Dynamic Panel Models: An Empirical Analysis of Cross-Country Data for the Study of Economic Growth', in C. Hsiao, K. Lahiri, L.-F. Lee and M. Pesaran (Eds.), *Analysis of Panel Data and Limited Dependent Variables,* Cambridge: Cambridge University Press.

Nerlove, M. (2000) 'Growth Rate Convergence, Fact or Artifact?', in Krishnakumar, J. and Ronchetti, E. (Eds.), *Panel Data Econometrics: Future Directions,* Amsterdam, North Holland.

Neumayer, E. (2002) 'Is Good Governance Rewarded? A Cross-National Analysis of Debt Forgiveness', *World Development,* 30(6), 913-930.

Niho, Y. (1974) 'Population Growth, Agricultural Capital, and the Development of a Dual Economy', *The American Economic Review,* 64(6), 1077-1085.

Nugent, J. (1996) 'What Explains the Trend Reversal in the Size Distribution of Korean Manufacturing Establishments?,' *Journal of Development Economics,* 48, 225-251.

Pagano, M., Panetta, F. and Zingales, L. (1998) 'Why Do Companies Go Public? An Empirical Analysis', *Journal of Finance,* 53, 27-64.

Pagano, P. and Schivardi, F. (2001), *Firm Size Distribution and Growth*, Banca d'Italia Working Paper 394.

Pakes, A. and Ericson, R. (1987), *Empirical Implications of Alternative Methods of Firm Dynamics*, mimeo, University of Wisconsin-Madison.

Petersen, M. and Rajan, R. (1995) 'The Effect of Credit Market Competition on Lending Relationships', *Quarterly Journal of Economics,* 110(2), 407-443.

Piore, M. and Sabel, C. (1984), *The Second Industrial Divide: Prospects for Prosperity*, New York, Basic Books.

Political Risk Services (1996), *International Country Risk Guide*. PRS Group, East Syracuse, New York.

Putterman, L. (2006), *Agricultural Transition Year,* Brown University, Providence, Rhode Island.

Pyke, F. et al. (1990), *Industrial Districts and Inter-firm Co-operation in Italy*, International Institute for Labor Studies, Geneva.

Quinn, D. (1997) 'The Correlates of Change in International Financial Regulation', *American Political Science Review,* 91(3), 531-551.

Radelet, S. and Sachs, J. (1998) 'Shipping Costs, Manufactured Exports, and Economic Growth', a paper presented at the 1998 American Economics Association annual meeting in Chicago. A copy of the paper can be obtained via internet at *http://www.earthinstitute.columbia.edu/about/director/pubs/shipcost.pdf*

Raftery, A. (1995) 'Bayesian Model Selection in Social Research', *Sociological Methodology,* 25, 111-163.

Rajan, R. and Zingales, L. (1998) 'Financial Dependence and Growth', *American Economic Review,* 88, 559-587.

Ranis, G. and Fei, J. (1961) 'A Theory of Economic Development', *The American Economic Review*, 51(4), 533-565.

Redding, S. and Venables, A.J. (2003), *Geography and Export Performance: External Market Access and Internal Supply Capacity,* NBER Working Paper 9637, National Bureau of Economic Research, Cambridge.

Renolds, L.G. (1969) 'Economic Development with Surplus Labor: Some Complications', *Oxford Economic Papers,* 21(1), 89-103.

Ringleb, A. and Wiggins, S. (1990) 'Liability and Large-Scale, Long-Term Hazards', *Journal of Political Economy*, 98(3), 574-595.

Rodrik, D., Subramanian, A. and Trebbi, F. (2002), *Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development,* National Bureau of Economic Research Working Paper 9305, National Bureau of Economic Research, Cambridge.

Rosen, S. (1982) 'Authority, Control and Distribution of Earning', *Bell Journal of Economics, Autumn, 311-323.*

Ryan, J. (1994) 'Survey Methodology: The Tabulated Ratings', in Freedom House, *Freedom in the World 1993-94.* New York.

Sachs, J.D. (2003), *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income,* NBER Working Paper 9490, National Bureau of Economic Research, Cambridge.

Sachs, J.D. (2003) 'Institutions Matter, but Not for Everything,' *Finance and Development,* 40(2).

Sala-i-Martin, X., Doppelhofer, G. and Miller, R. (2004) 'Determinants of Long-Run Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach', *American Economic Review*,94, 813-835.

Sanchez-Robles, B. (1998) 'Infrastructure Investment and Growth: Some Empirical Evidence', *Contemporary Economic Policy*, 16(1), 98-108.

Sato, R. and Niho, Y. (1971) 'Population Growth and the Development of a Dual Economy', *Oxford Economic Papers,* 23(3), 418-436.

Scherer, F. (1980), *Industrial Market Structure and Economic Performance,* Boston, Houghton Mifflin.

Schivadi, F. and Torrini, R. (2003), *Firm Size Distribution and ELP in Italy,* William Davidson Institute Working Paper 613, The William Davidson Institute,

Michigan.

Schivadi, F. and Torrini, R. (2004), *Threshold Effects and Firm Size: The Case of Firing Costs,* CEP Discussion Paper 633, Center for Economic Performance, London.

Schultz, T.W. (1964), *Transforming Traditional Agriculture,* Chicago: The University of Chicago Press.

Schumpeter, J.A. (1911), *The Theory of Economic Development*, Cambridge, MA: Harvard University Press.

Schwarz, G. (1978) 'Estimating the Dimension of a Model', *Annals of Statistics,* 6, 461-464.

Sen, A.K. (1966) 'Peasants and Dualism with or without Surplus Labor', *The Journal of Political Economy,* 74(5), 425-450.

Sengenberger, W., Loveman, G. and Piore, M. (1990), *The Re-Emergence of Small Enterprises: Industrial Restructuring in Industrialized Countries,* International Labor Office, International Institute for Labor Studies, Geneva.

Sengenberger, W. and Pyke, F. (1990), Small *Firm Industrial and Local Economic Regeneration: Research and Policy Issues,* International Institute for Labor Studies, Geneva.

Shaw, E. (1973), *Financial Deepening in Economic Development,* New York: Oxford University Press.

Shea, J. (1997) 'Instrument Relevance in Multivariate Linear Models: A Simple Measure', *Review of Economics and Statistics,* 79(2), 348-352.

Simon, H. and Bonini, C. (1958) 'The Size Distribution of Business Firms', *American Economic Review,* 48, 607-617.

Sleuwaegen, L. and Goedhuys, M. (2002) 'Growth of Firms in Developing Countries, Evidence from Cote d'Ivoire', *Journal of Development Economics,* 68, 117-135.

Snodgrass, D. and Biggs, T. (1996), *Industrialization and the Small Firm. Patterns and Policies,* International Center for Economic Growth and Harvard Institute for International Development, San Francisco, California.

Spilling, O. (1998) 'On the Re-emergence of Small Scale Production: The Norwegian Case in International Comparison,' *Small Business Economics,* 10, 401-417.

Stock, J., Wright, J. and Yogo, M. (2002) 'A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments,' *Journal of Business and Economic Statistics,* 20(4), 518-529.

Storey, D. (1994), *Understanding the Small Business Sector,* London: Routledge.

Sutton, J. (1997) 'Gibrat's Legacy', *Journal of Economic Literature,* 35(1), 40-59.

Tambunan, T. (2005), *Rethinking Small and Medium Enterprise as a Growth Engine in Developing Economies,* University of Krisakti.

Taymaz, E. and Kaliçaslan, Y. (2005) 'Determinants of Subcontracting and Regional Development: An Empirical Study on Turkish Textile and Engineering Industries', *Regional Studies,* 39, 633-645.

Temple, J. (2000) 'Growth Regressions and What the Textbooks Don't Tell You', *Bulletin of Economic Research,* 52, 181-205.

Temple, J. and Woßmann, L. (2006) 'Dualism and Cross-Country Growth Regressions', *Journal of Economic Growth,* 11, 187-228.

Thomasdakis, S. and Droucopolous, V. (1996) 'Dynamic Effects in Greek Manufacturing: The Changing Share of SMEs, 1983-1990', *Review of Industrial Organization,* 11, 69-78.

Thornburg, L. (1993) 'IBM Agent's of Influence', *Human Resource Magazine,* 38(2).

Todaro, M.P. and Smith, S.C. (2003), *Economic Development,* Essex: Pearson Education Limited.

Trau, F. (1997) 'Recent Trends in the Size Structure of Italian Manufacturing Firms', *Small Business Economics,* 9, 273-385.

Uanto, G. and Sanchez, M. (1998), *Financial Liberalization and Small-Scale and Micro Enterprises in the Philippine Manufacturing Industry,* MIMAP Working Paper No. 38, MIMAP, Manila, Philippines.

UNIDO (2006), *Industrial Statistics Database,* United Nations Industrial Development Organization, Vienna, Austria.

Viner, J. (1932) 'Cost Curves and Supply Curves', *Zeitschnft fur National-Skonomie,* 3, 23-46.

Wagner, J. (1995) 'Exports, Firm Size, and Firm Dynamics,' *Small Business Economics,* 7(1), 29-39.

Wagner, J. (2001) 'A Note on the Firm Size-Export Relationship,' *Small Business Economics,* 17, 229-237.

Weber, E. (1976), *Peasants into Frenchmen,* California: Stanford University Press.

Weeks, J. (2003), 'Small Manufacturing Establishments in Developing Countries', *International Review of Applied Economics*, 17(4), 339-359.

Williamson, O. (1975), *Market and Hierarchies: Analysis and Antitrust Implications*, New York: The Free Press.

Williamson, O. (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, New York: The Free Press.

World Bank (1994), *Can Intervention Work? The Role of Government in SME Success*, World Bank, Washington, D.C.

World Bank. (2002), *SME, World Bank Group Review of Small Business Activities*, World Bank, Washington, D.C.

World Bank. (2004), *SME, World Bank Group Review of Small Business Activities*, World Bank, Washington, D.C.

World Bank, *Doing Business Database*, World Bank Group, Washington, D.C.

You, J-I. (1995) 'Small Firms in Economic Theory: Critical Survey', *Cambridge Journal of Economics*, 19, 441-462.