DISSECTING THE SIZE DISTRIBUTION OF ESTABLISHMENTS ACROSS COUNTRIES

Manuel Garcia-Santana and Roberto Ramos

CEMFI Working Paper No. 1204

May 2012

CEMFI
Casado del Alisal 5; 28014 Madrid
Tel. (34) 914 290 551 Fax (34) 914 291 056
Internet: www.cemfi.es

We thank seminar participants at CEMFI for very useful comments. Manuel and Roberto gratefully acknowledge financial support of the Spanish Ministry of Education under research grant BES-2009-025673 and BES-2009-026803, respectively.

CEMFI Working Paper 1204
May 2012

DISSECTING THE SIZE DISTRIBUTION OF ESTABLISHMENTS ACROSS COUNTRIES

Abstract

We present a detailed description of cross-country differences in the size distribution of establishments as measured by Enterprise Surveys of the World Bank (2006-2010). We find that poorer countries tend to have smaller establishments, and hence a higher proportion of employment is allocated to such plants. We also find that, conditional on the level of per capita GDP, there is still a wide variation in the size distribution of establishments. We show that financial frictions and other costs imposed on the business environment can account for a sizeable part of such variation. Additionally, we exploit the richness of the data-set to document cross-sector differences in the size distribution of establishments. We show that establishments in manufacturing and construction tend to be larger on average than in retail and other services.

Keywords: Plant size distribution, enterprise environment.

JEL Codes: L11, L53, O47.

Manuel Garcia-Santana
CEMFI
m.g.santana@cemfi.edu.es

Roberto Ramos
CEMFI
roberto.ramos@cemfi.es

1 Introduction

Why do some countries have so low levels of income per capita? Why, for instance, income per capita in Nepal is only 2 percent that of the United States? A common view is that a high proportion of income variation across countries can be attributed to differences in total factor productivity (TFP).[1] Moreover, a recent strand of literature has started to emphasize misallocation of resources across plants as a source of these differences in aggregate productivity.[2] This misallocation takes place due to the existence of government policies that distort the way in which resources are allocated across heterogeneous production units. This implies that the size distribution of plants is a crucial object in order to understand the income differences across countries, because it reflects how efficiently limited resources are allocated.

However, there has not been much attention to studying empirically the cross-country variation in size distribution of plants. In old works, Banarji (1978) shows for a small number of countries that the average size of plants is positively correlated with physical capital intensity. Liedholm and Mead (1987) provide evidence of poor countries having most of the employment allocated to small and large plants, establishing a phenomenon known as 'the missing middle'.[3] In a classic paper, Tybout (2000) collects all this previous evidence to discuss the poor performance of the manufacturing sector in developing countries. Leaning on country-level studies, he argues that a strong business regulation could be behind the excessive presence of small entrepreneurs. Remaining small, entrepreneurs avoid government regulation and hence do not have incentives to achieve a larger size.[4] In a more recent work, Alfaro et al. (2008) use establishment level data for 79 countries to calibrate a Melitz (2003) type model in order to infer the level of distortions necessary to generate the deviation in the distribution of establishments with respect to the one in the US. In a close paper to our work, Poschke (2011) documents that the average, standard deviation, and skewness of the size distribution of firms are positively correlated to income per capita, using firm-level data for around 50 countries. He also finds that the entrepreneurship rate is

[1] See Caselli (2002), Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999).

[2] Restuccia and Rogerson (2008), Guner et al. (2008) and Hsieh and Klenow (2009) are some examples

[3] In our work, the size of a plant is defined in terms of number of employees.

[4] See De Soto (1989) for Peru, Mohan (2002) for India and Lewis (2005) for a small set of countries

negatively correlated to income per capita. He then calibrates a version of Lucas (1978) that is able to account for these four facts.

Our paper contributes to the literature mentioned above, providing additional evidence about cross-country differences in how resources are allocated across heterogeneous production units. We use the Enterprise Surveys of the World Bank (ESWB) for the period 2006-2010. This dataset is specially suitable to analyze the size distribution of plants across countries. This is mainly for three reasons. First, it is standardized. This means that every plant in every country answers the same questions and therefore cross-country comparisons are possible. Second, the sample of surveyed plants is representative of the population of formal private non-agricultural plants. This allows us to establish some facts about the allocation of resources beyond manufacturing. And third, coverage is very broad. We use data on 128 country-surveys, which give us power to validate the statistical significance of our findings. Moreover, surveyed countries are mostly of low and middle income per capita, hence, most likely to be affected by distortions that disturb the allocation of resources.

We start showing that small plants are associated to lower levels of productivity relative to big plants. This underlines the importance of studying the size distribution of establishments to understand cross-country income differences. Then, we analyze the size distribution of the 128 country-surveys in our sample, disaggregated by sector. We find huge variation in the size distribution across countries, as shown by Figure I. We provide evidence of the following facts. To the best of our knowledge, facts (iii), (iv), (v) and (vi) have not been documented before.

(i) The average establishment size is lower in poor countries than in relatively richer countries. This is true not only in the manufacturing sector but also in services.

(ii) Consistent with fact (i), the share of employment allocated to small plants in poor countries is larger than in higher income countries. Excluding foreign owned firms, this fact is quantitatively more important.

(iii) Conditional on per capita GDP, the share of employment accounted by small plants is positively correlated with the level of distortions in the economy.

(iv) Also conditional on per capita GDP, financial frictions in particular are strongly positively correlated with how much labor is allocated to small plants. This

association is magnified in retail and wholesale sector.

(v) Controlling for per capita GDP and the level of distortions, larger countries tend to allocate a higher amount of labor to large plants.

As mentioned above, papers such as Banarji (1978) and Liedholm and Mead (1987) already document facts (i) and (ii) for the manufacturing sector. Also, Poschke (2011) finds the same pattern using more recent data and a larger set of countries. We extend these results exploiting the broadest country coverage of ESWB, that allows us to disaggregate the size distribution at the sectoral level.

We stress facts (iii) and (iv): conditional on per capita income, we still find wide variation in the size distribution of plants. In other words, we document a wide within income-levels variation in how resources are allocated across countries. This disparity is specially high at low levels of income, suggesting that in poor countries there is high variation in the type of policies and other distortions that affect the allocation of resources. Additionally, we find that the business environment and access to credit explain a sizable part of this within-income variation. In particular, we find that countries with higher levels of distortions tend to allocate more resources to small plants. This is significant, because controlling for per capita income we keep fixed potential determinants of the size distribution of plants other than distortions.

These results are consistent with the cross-country implications of a recent influential literature that uses theoretical frameworks to quantitatively measure the marginal effects of the presence of distortions. This literature shows that the existence of distortions prevents an optimal allocation of resources. In particular, distortions make too many resources being allocated to small unproductive firms, generating a high efficiency loss and hence creating big output losses. Guner et al. (2008) show that policies that reduce the average size of establishments by 20 per cent lead to reductions in output up to around 8 per cent. Hsieh and Klenow (2009) find that removing distortions in India and China such that marginal products are equalized to the extent observed in the US would imply TFP gains of up to 50 per cent in China and up to 60 per cent in India.

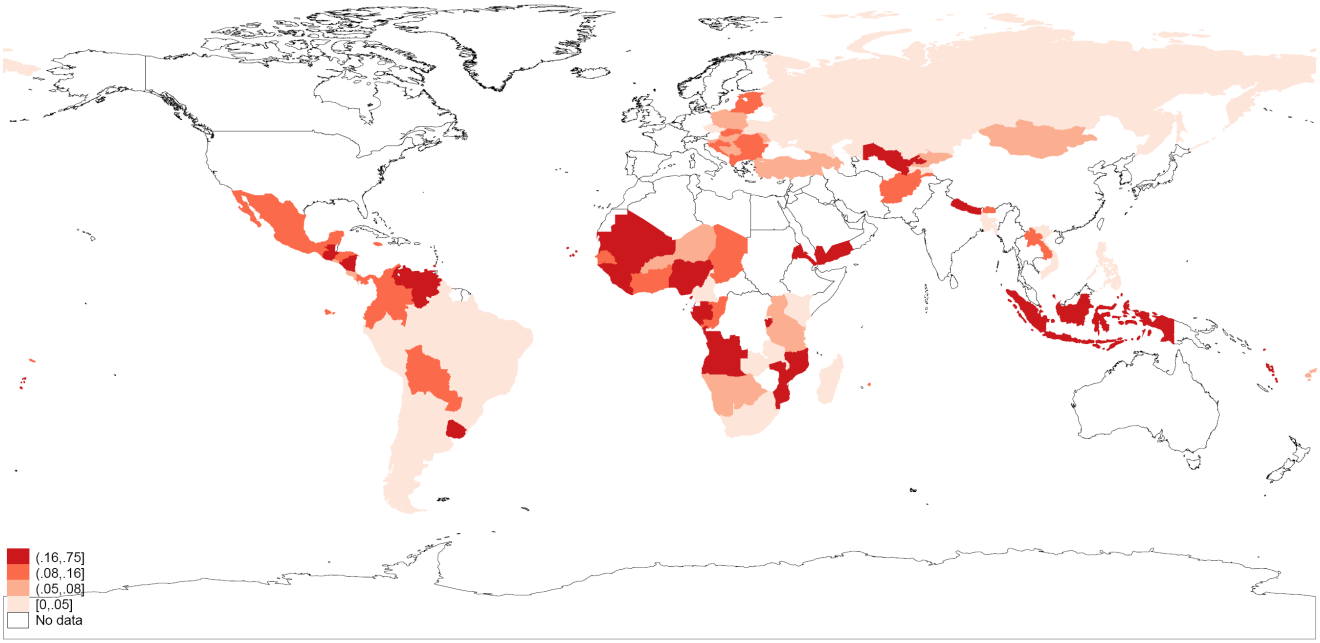
Finally, we explore the existence of the phenomenon known as the 'missing middle' i.e. the relative absence of firms of medium size, as compared to small and large firms, first documented by Liedholm and Mead (1987). We propose a simple methodology to

uncover this phenomenon and use it to comprehensive data on Indian firms. Using our dataset of ESWB, we are unable to systematically analyze the missing middle in our sample of countries. This is mostly due to the fact the micro-entreprises are left out of the survey, a limitation that is pervasive in the size distribution literature.

The rest of the paper is organized as follows. In Section 2 we explain in detail the characteristics of our dataset and compare it with other databases used to study the size distribution of plants. In Section 3 we show that small plants tend to be less productive than medium and big plants. In section 4 we document the basic facts presented in this paper, underlining the cross-country correlations between income per capita and the importance of small plants. In section 5 we present our econometric framework to study the associations between the size distribution of plants and distortions. In section 6 we study the phenomenon known as 'the missing middle'. Section 7 concludes.

FIGURE I

SHARE OF LABOR ACCOUNTED BY SMALL PLANTS
MANUFACTURING (2006-2010)



2 Enterprise Surveys of the World Bank

In order to analyze the size distribution of firms for a large sample of countries, we use Enterprise Surveys of the World Bank (ESWB). These are an impressive collection of plant-level surveys meant to be representative of a country's non-agricultural private formal economy. The goal of these surveys is to collect information about the business environment and how it affects the performance of plants across developing countries.

Originally started in 2002, we use the Standardized Data 2006-2010.[5] This dataset has a number of advantages that makes it unique and very helpful to dissect the size distribution of establishment across countries. First, it provides standardized micro data across a very broad range of developing countries. This means that all firms in all countries are asked the same questions. Therefore, it provides a crucial advantage for making reliable cross-country comparisons.

Second, interviewed firms in each country are carefully meant to be representative of the targeted population. The sampling methodology is stratified random sampling with replacement. That is, homogeneous groups are selected and random sampling is performed on each group.[6] To perform population estimates, therefore, properly weighting is necessary. The standardized dataset 2006-2010 provides weights for all the observations.[7]

And third, coverage is very broad. Our sample consists of 128 surveys corresponding to 104 countries, most of them of low and middle income per capita, according to the classification of the World Bank.[8] Therefore, the sample is biased towards countries in the earliest stages of development. Per capita GDP of the median country in our sample -Tonga- is 10 per cent that of the US. Income of the first quartile -Costa Rica- is slightly less than one quarter of US per capita income. This allows us to study changes in the size distribution as countries move forward in the ladder of development.

The Enterprise Surveys are answered by business owners and top managers.[9] Typ-

[5] Available at https://www.enterprisesurveys.org/.

[6] In the case of the ESWB, strata are based on firm size, business sector and geographic region within a country.

[7] A Standardized Dataset for the period 2002-2005 is also available. However, just 35 per cent of the observations have information on weights. Hence, it is not possible to estimate unbiased population statistics for the majority of countries

[8] See Table B.1 for a list of all the countries included in the sample.

[9] Surveys are conducted at the establishment level, although some information of the firm is asked at the beginning of the questionnaire, like ownership. Although we make clear this point, given that

ically 1200/1800 interviews are conducted in large economies, 360 in medium-sized economies and 150 in small economies. Addressed topics include firm characteristics, workforce, sales, ownership, corruption, finance and obstacles to growth, among others. Crucially, information is provided on the business sector, according to the classification ISIC rev 3.1. We use this information to compute the size distribution at the country-sector level.

Two limitations of the dataset are worth mentioning. First, firms with less than 5 employees are not targeted. And second, only the formal economy is covered. Since firms in the informal economy are more likely to be small, both issues lead our measures of the size distribution to be biased towards a lower amount of employment allocated to small firms.[10]

The first shortcoming is a common one in the literature of heterogeneous firms. For instance, Alfaro et al. (2008) truncate the data in 20 employees, as countries with low coverage in their database are very likely to overrepresent older and larger establishments. Also, Hsieh and Klenow (2009) work with Indian plants of more than 10 workers and with non-state Chinese firms of more than 5 million yuan in revenue. Regarding the second limitation, we control in our regressions for the size of the informal sector. These two shortcomings are not present in the data-set used by Poschke (2011), whose main advantage is that it captures any kind of firm, regardless of its size or legal status. The reason is that the survey is carried out at the individual/entrepreneur level, which makes it very attractive to study occupational choice.[11] However, the sample of countries is small relative to ESWB. Table I shows the main characteristics of data sets used for cross-country analysis of size distribution of production units.

In our sample, one quarter of the countries are low-income economies and only 8 are OECD member countries. This raises the natural concern of how accurate is the measurement of employment in the dataset, as surveys are in principle less reliable for

85 per cent of the observations are firms of one establishment, in this paper we use the words firm, establishment and plant interchangeably.

[10]Additional surveys covering informal enterprises -Informal Surveys- and firms with less than 5 employees -Micro Surveys- have been developed by the World Bank. However, they have not been integrated to the standardized data yet. For further information on ESWB, the notes 'Understanding the Sampling Methodology', 'Understanding the Questionnaire' and the 'Implementation Notes', available online, are very useful to comprehend the rationale of the sampling, the variables contained in the dataset and the specifics of each country-survey, respectively.

[11]The dataset is Global Entrepreneurship Monitor, available at http://www.gemconsortium.org/

TABLE I

COMPARISON BETWEEN DIFFERENT DATA-SETS

Paper	Name	Countries	Level of survey	Truncation	Informal
Alfaro et al. (2008)	Dun & Bradstreet	80	Plant	It varies	No
Poschke (2011)	GEM	50	Firm	0	Yes
This	ESWB	104	Plant	5	No

poorer countries. To tackle this concern, we compare the total number of workers estimated from the ESWB to the total number of workers provided by an alternative data source, the Penn World Table 7.0 (PWT) -Heston et al. (2011)-. Note that stratified random sampling with appropiate weights allows us to compute population estimates from the ESWB data.

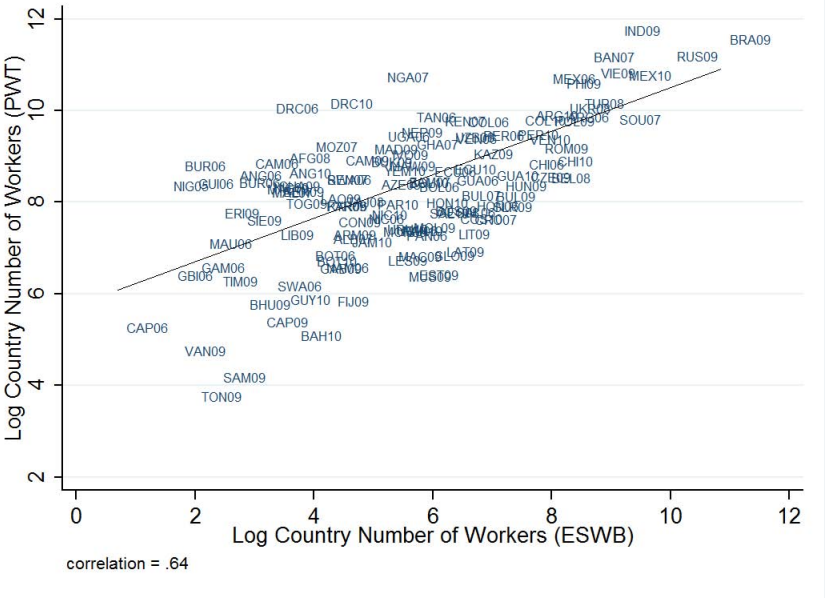
For each country, we estimate the total number of workers in the sectors targeted by the Enterprise Surveys weighting appropiately each observation. Panel A of Figure VII shows the correlation between this estimate and the country total number of workers calculated from the PWT. Both variables are highly correlated (.64), suggesting a fair degree of accurateness of the Enterprise Surveys data.

Still, there exists some dispersion between the number of workers calculated from both data sources. Note however that both calculations are not strictly comparable. As already mentioned, ESWB do not target all the economic sectors. Importantly, firms in agriculture and informal enterprises are not targeted, as well as state-owned companies. On the contrary, PWT takes into account overall employment including agriculture, the shadow economy and state-owned companies. Therefore, to make a more reliable comparison, we regress the log number of workers provided by the PWT against the log number of workers calculated from the ESWB, controlling for the share of employment in agriculture and the size of the informal economy.[12]

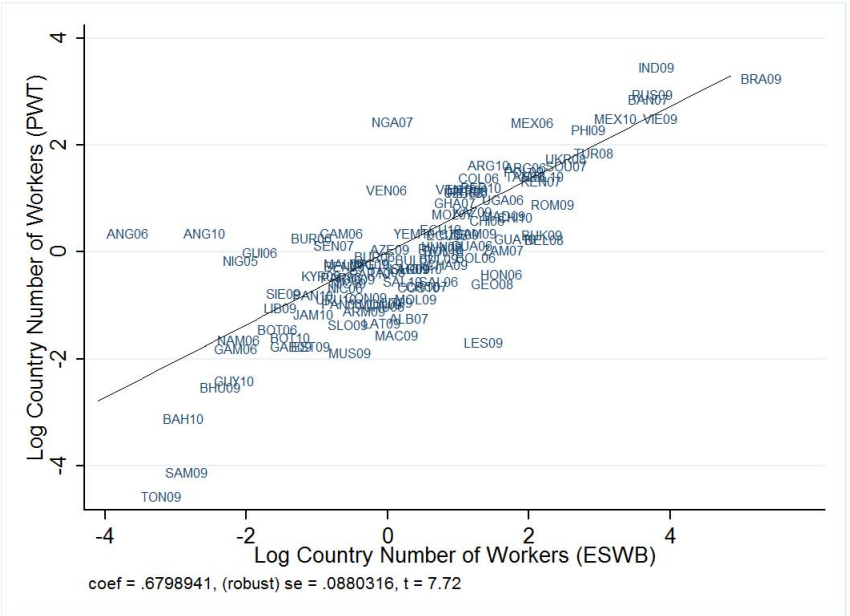
Panel B of Figure VII shows the partial correlation between the number of workers calculated from the ESWB and that reported by the PWT, controlling for employment in agriculture and the shadow economy. As is evident, the dispersion is significantly reduced. Some differences remain for some countries though, specially at the lower end of the employment distribution, but overall this rough comparison is very supportive on the quality and standards of the Enterprise Surveys, at least when dealing with

[12]See Appendix A for a definition of all variables and sources used throughout the paper.

Figure II
Correlation between ESWB and PWT



(a) Correlation



(b) Partial Correlation

employment.[13]

3 Size and Productivity at the plant-level

As mentioned in the Introduction, our working hypothesis is that the allocation of factors across heterogeneous production units matters in trying to understand income differences across countries. In particular, we work under the hypothesis that a higher amount of resources allocated to small plants is associated to lower levels of aggregate productivity. In this section we provide some broad descriptive evidence at the plant level to show that small plants tend to be less productive than large plants. To this end, we run the following regression:

$$\begin{aligned} log\frac{VA_{ija}}{N_{ija}} &= \gamma_0 + \gamma_1 \text{Size}_i + \beta_2 \text{logAge}_i + \beta_3 \text{Foreign}_i \\ &\quad + \beta_4 \text{ExportStat}_i + \sum_j \mu_j + \sum_a \nu_a + u_{ija} \end{aligned} \tag{1}$$

where $\frac{VA_{ija}}{N_{ija}}$ is valued added divided per employee of plant i in sector j and country a; $Size$ can be whether plant i is small, medium or large or the log number of employees of the plant; $logAge$ is the log of the number of years during which the plant i has been operating; $Foreign$ is a dummy that takes value one of the plant i is foreign owned and $ExportStat$ is a dummy that takes value one if the plant i exports and zero otherwise. We also include sector and country dummies.

Table II shows the results under different specifications. All columns represent regression results where the log value added per employee is the dependent variable and standard errors are clustered at the country-survey level. In columns (1) and (2) the definition of size is discrete. That is, we include dummies for medium and large plants, being the 'small' status the ommitted one. In columns (3) and (4) size is defined as the log number of employees.

The coefficients associated to size are positive and significant in all four specifications. In fact, the value of these coefficients is very high. For instance, in column (2) we

[13]Dropping those countries for which these differences are the highest increase the estimates found in all the sections.

see that conditional on country, export status, foreign ownership, age, and sector of operation, large plants are associated to around a 80 per cent higher level of productivity. We also find that, everything else equal, foreign and exporters plants tend to be more productive. For instance, in column (4) we see that foreign ownership is associated to around a 30 level of productivity.

Table II
Firm-Level Correlation between Size and Productivity

	(1)	(2)	(3)	(4)
Dep. Variable: Log Productivity				
Medium Establishment	0.3838***	0.2731***		
	(0.0614)	(0.0700)		
Large Establishment	1.0309***	0.8336***		
	(0.2048)	(0.2458)		
Log Employees			0.3223***	0.2466***
			(0.0772)	(0.0769)
Foreign-Owned		0.3397**		0.3226**
		(0.1456)		(0.1302)
Export Status		0.2579***		0.2225***
		(0.0431)		(0.0554)
Log Age		0.0446		0.0354
		(0.0307)		(0.0273)
Sector Dummies	SECTOR	ISIC	SECTOR	ISIC
Country Dummies	YES	YES	YES	YES
Observations	26,878	26,283	26,878	26,283
R-squared	0.80	0.84	0.81	0.84

Table II shows firm-level regressions of log Productivity against Size of the establishment. Columns (1) and (2) computes size under three categories -small (excluded), medium and large establishments-, defined as less than 20 employees, between 20 and 100 employees and more than 100 employees. Columns (3) and (4) computes size as the log number of employees. Sector dummies are either SECTOR -Manufactuing, Construction, Trade and Services- or ISIC Rev 3.1 at the two-digit level. Standard errors are in parenthesis, clustered at the country-survey level. Significance levels: *: 10%; **: 5%; ***: 1%. See Appendix A and the text for further details.

4 Size Distribution of Establishment Across Countries: A Broad View

We use the Standardized Data 2006-2010 of the World Bank to construct the size distribution of establishments across the 128 country-year observations in our sample. For each establishment, we compute size as the number of permanent plus temporary workers at the end of last fiscal year. We consider four different sectors per country: manufacturing, construction, trade (retail and wholesale) and other services.

We start our characterization of the size distribution with an analysis of the average size of establishments across countries and sectors. We first group countries according to their level of income. We follow the classification of the World Bank, who categorizes countries according to the Gross National Income per Capita in 2010. The groups are: low income, $1,005 or less; lower middle income, $1,006 - $3,975; upper middle income, $3,976 - $12,275; and high income, $12,276 or more. As in our sample we have only 8 high income countries, we include them in the group of upper middle income. We then calculate average log size of each group of countries for each sector.

Table III shows the results. We observe that in all sectors the average size of establishments is smaller in low income countries, followed up by low middle income economies and then by upper middle & high income countries. The only exception is construction, where this ranking is not preserved. Moreover, in column (4) we perform a test of these differences in the average log size being statistically significantly different from zero. We compare the group of countries with the relatively highest income with that of the lowest income per capita -columns (3) and (1), respectively-. In all cases except construction, we reject the null hypothesis of the difference being statistically equal to zero.[14]

The difference in average log size between relatively richer and poorer countries is highest in trade, where the average plant is approximately 35 per cent larger in the former as compared to the latter. In services this difference is approximately 25 per cent, whereas in manufacturing is 20 per cent and in construction 9 per cent. This

[14]We have also performed a regression of the average log size of the country-sector against per capita GDP relative to the US and including sector fixed effects. The coefficient on GDP per capita is .47, significant at the 90 per cent confidence level. If we exclude construction, which represents less than 6 per cent of the economic activity in our data, the coefficient increases to .59, and is significant at the 95 confidence level.

Table III

Average Log Establishment Size across Groups of Countries and Sectors

	Low Income Countries	Low Middle Income Countries	Upper Middle & High Income Countries	T-test
	(1)	(2)	(3)	(4)
Manufacturing	3.07	3.12	3.27	0.0930
	(0.57)	(0.47)	(0.34)	
Construction	3.40	3.39	3.49	0.6360
	(0.82)	(0.68)	(0.56)	
Services	2.70	2.80	2.94	0.0228
	(0.41)	(0.46)	(0.48)	
Trade	2.37	2.62	2.72	0.0000
	(0.31)	(0.39)	(0.35)	

Table III shows the average log size of establishments across countries -classified according to their level of income- and sectors. Standard deviations in parenthesis. The fourth column displays the p-values of a test with null hypothesis being that the average of the upper middle & high income countries is different from the average of the low income countries, allowing for unequal variances. See Appendix A and the text for further details.

suggests that those elements that the literature has emphasized as determinants of the misallocation of resources in poor countries are more prevalent in trade and services than in manufacturing. We come back to this question in Section 5.[15]

Next, in order to deeper characterize the size distribution of establishments across countries, we compute, for each country and each sector, the labor allocated to small -less than 20 employees-, medium -between 20 and 99 employees- and large plants -100 or more employees-. That is, for each sector j and country a, we compute the size distribution of establishments as:

$$S^s_{a,j} = \frac{\sum_{i=1}^{N} L_{i,a,j} \mathbf{1}\{L_{i,a,j} \leq 19\}}{\sum_{i=1}^{N} L_{i,a,j}} \tag{2}$$

[15]The finding that average log size of establishments is increasing with per capita GDP contrasts with Alfaro et al. (2008), who find the opposite result. Most likely, the reason for such disparity lays in the sample of countries considered. Our sample is comprised mostly by developing countries, with average per capita GDP being 14 per cent that of the US. In Alfaro et al. (2008), the sample considered includes 24 OECD member countries -out of 79 countries- and average per capita GDP is 32 per cent that of the US, more than twice as much as in our sample. Moreover, Alfaro et al. (2008) truncate the distribution of employment below 20 employees. Anyhow, this disparity is worth exploring. We leave it for further research.

$$S^m_{a,j} = \frac{\sum_{i=1}^{N} L_{i,a,j} \mathbf{1}\{L_{i,a,j} \in [20, 99]\}}{\sum_{i=1}^{N} L_{i,a,j}} \tag{3}$$

$$S^l_{a,j} = \frac{\sum_{i=1}^{N} L_{i,a,j} \mathbf{1}\{L_{i,a,j} \geq 100\}}{\sum_{i=1}^{N} L_{i,a,j}} \tag{4}$$

where $S^s_{a,j}$ $S^m_{a,j}$ and $S^l_{a,j}$ are the amount of labor allocated to small, medium and large establishments, respectively. These categories of size correspond to the ones defined by the World Bank. As before, we group countries according to their income level.

Table IV shows the results. For all sectors, the amount of labor allocated to small establishments is decreasing in the level of per capita income. For instance, in manufacturing, on average, 18 per cent of the total work force is allocated to small establishments in low income countries, whereas this percentage is 13 per cent and 9 per cent in low middle and upper middle & high income economies, respectively. The amount of labor allocated to medium plants is lower in countries with upper middle and & high income per capita. However, no difference arises between low income and low middle income countries.

As a result, the pattern observed in the amount of labor allocated to small plants is reversed when considering the workforce employed in large establishments. On average, in manufacturing, low income economies employ 54 per cent of the workforce in small plants; low middle income economies employ 59 per cent; and upper middle & high income countries employ 65 per cent. Again, this rising pattern is constant across sectors.

Within sectors, we observe in general that manufacturing and construction tend to employ a higher proportion of the labor force in big plants than do trade and other services. Richer countries concentrate more than 60 per cent of the workforce in large plants in manufacturing and construction, whereas just slightly above 50 per cent in trade and services. This cross-sectoral pattern is more severe in the poorest countries. For instance, large firms in manufacturing employ 54 per cent of the workers in low income economies. In trade, on the contrary, large firms employ just half as much. Actually, for countries with income equal or below the low middle threshold, services and trade allocate the bulk of employment to small and medium establishments. This is an important finding because it suggests that poor countries perform relatively better in manufacturing than in services.

Summing up, looking at the size distribution of establishments we observe three patterns. First, average size of establishment is higher the higher is per capita income of the country. This is true not only in manufacturing but also in trade and services. Second, regardless of the sector, poorer countries tend to allocate a higher proportion of the labor force in small plants, and a lower proportion in big plants.[16] And third, regardless of the income level, manufacturing and construction employ a higher amount of the workforce in large establishments than do trade and services, whereas the reverse is true for small establishments. This pattern is specially acute in countries with low income per capita.

As emphasized by Garcia-Santana (2011), an important determinant of the size distribution of plants across the developing world is the presence of foreign firms. These tend to be larger on average than domestic firms and, as modeled by Garcia-Santana (2011), its presence may generate a crowding out effect on unproductive domestic entrepreneurs. These two effects go in the same direction, resulting in a higher share of labor explained by small firms in countries where the presence of foreign firms is lower.

We investigate this issue analyzing the size distribution of domestic plants, which is shown in Table VII. That is, we exclude those establishments that belong to firms that are owned by foreign agents. We define a firm as foreign owned if private foreign individuals, companies or organizations own 50 per cent of more of the firm.[17]

A comparison of Tables IV and VII shows that, once we exclude foreign owned firms, a higher amount of labor is employed in small firms. This is true in all sectors and in all income levels. The workforce employed in domestic firms is also more likely to be allocated to plants of medium size than when considering all types of firms. The exception is services and trade in low income countries, in which medium domestic firms explained a lower or equal share of labor than that explained by domestic and foreign owned companies. Regarding large firms, excluding foreign firms decreases the share of labor allocated to them. This is true in all sectors and for all income levels.

Importantly, excluding foreign owned establishments implies that the differences in the allocation of labor between richer and poorer countries are exacerbated. Note, for

[16]For all sectors, these differences between low income and upper middle & high income countries are significant under a t test of equality of means, allowing for unequal variances.

[17]Most firms with some foreign capital are 100 per cent foreign owned, so the results are not sensitive to enlarging the definition of foreign ownership to firms with a strictly positive amount of foreign capital.

instance, the share of labor allocated to large establishments in manufacturing. The low income over upper middle & high income ratio of this variable is 0.83 when we consider both domestic and foreign plants. Excluding foreign plants reduces this ratio to 0.71. Note also that this pattern is present across all sectors. This finding is consistent with two different hypotheses presented in Burstein and Monge-Naranjo (2009) and Garcia-Santana (2011). In both papers, an efficiency gain of foreign investment takes place in the host countries from reallocating domestic resources from domestic to foreign firms. In the former, this is the case because foreign firms have a higher average firm-embedded productivity. In the latter, foreign firms have access to less distorted financial markets and hence, in contrast to the domestic ones, theys are able to produce at their optimal scale.

5 Size Distribution and Distortions

A recent strand of literature emphasizes misallocation of resources across firms as a source of aggregate income differences across countries. This literature shows that the existence of policies distorting the equilibrium size distribution of firms generates high efficiency losses. Restuccia and Rogerson (2008), Guner et al. (2008) and Hsieh and Klenow (2009) are examples of this line of research. Moreover, recent papers complement these findings, measuring the effects of particular policies in developing countries. See for instance Bollard et al. (2011) and Garcia-Santana and Pijoan-Mas (2011) for India, and Gallipoli and Goyette (2011) for Uganda.

A common result arises in these studies: higher levels of misallocation are associated to more resources being assigned to small plants. In this section, we provide a comprehensive cross-country analysis of this result. We study whether economies with higher levels of distortions in the business environment allocate more resources to small production units.

We use the *Ease of Doing Business Index* of the World Bank as our measure of distortions. Later, we complement the analysis focusing on an specific distortion that is well documented to have a sizeable impact on the allocation of resources, namely financial constraints.[18]

In all our analysis, we compare countries with the same level of income per capita.

[18]Some examples are Erosa and Hidalgo-Cabrillana (2008) and Buera et al. (2011)

We do so in order to control for potential determinants of the size distribution of plants other than distortions. For instance, the efficiency of available technologies is likely to be lower in poorer countries, which may have a direct effect on the size distribution. Also, the distribution of entrepreneurial talent may depend on the distribution of education in the population, which tends to be more skewed towards the left tail in poorer countries.

Size Distributions and Distortions: Doing Business Index

As mentioned above, in this section we study whether countries with a higher level of distortions are associated to a larger share of labor allocated to small establishments. To this end, we calculate the correlation of the amount of labor working in small plants and the *Doing Business Index*. This evaluates business regulations that enhance the economic activity and those that constraint it in 183 economies, ranking countries according to their business-friendly regulations, where a higher ranking means a worse business environment. Ten areas are covered: starting a business, dealing with construction permits, getting electricity, registering property, getting credit, protecting investors, paying taxes, trading across borders, enforcing contracts and resolving insolvency. Therefore, it is a very suitable indicator to summarize the distortions that can potentially have impact in the allocation of resources across plants.

Figure III shows the correlation between the *Ease of Doing Business Index* and real per capita GDP. Note that this index has support [1 183].[19] As it is patent, countries with a higher level of income exhibit a better business environment. The correlation between both variables is -0.58.

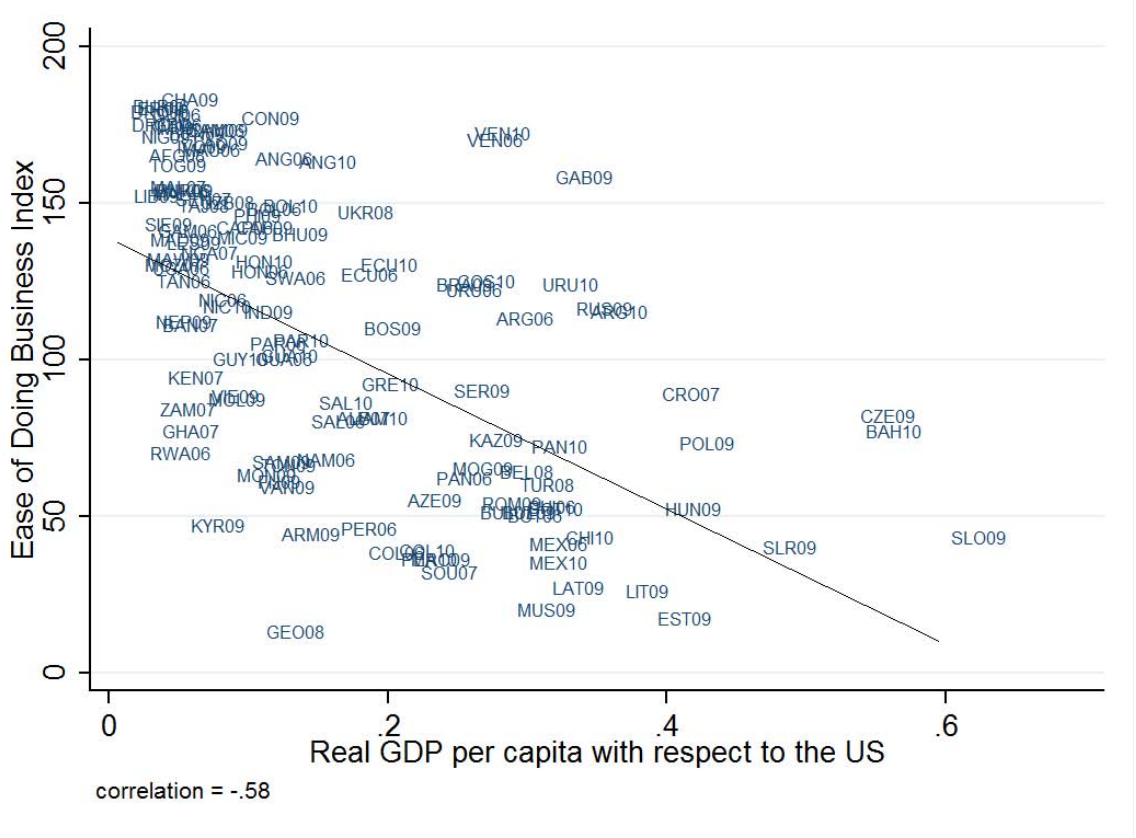
To analyze the explanatory power of distortions with respect to the size distribution, we regress the proportion of labor allocated to small plants on the *Doing Business Index*. We add three covariates, per capita GDP, the size of the informal economy and total size of the country. As already mentioned, we keep constant the countries' income level in order to control for other potential determinants of variation in the size distribution of plants, such as technology.

Total size of the country is captured by log total population. This aims to control for the size of the internal market and transportations costs within a country. Bigger markets may influence the optimal size of busineses, specially in trade and services,

[19] At the time of running regressions, we divide the support by 100 for presentational purposes

Figure III

Correlation between Ease of Doing Business Index and per Capita GDP



where goods are not tradable.[20]

Introducing the size of the informal economy allows us to tackle a possible ommitted-variable bias. The intuition is as follows: income per capita, the quality and the business environment and the size of the country determine the amount of labor working in small plants, both registered and unregistered. Therefore, conditional on such determinants, a higher presence of informal firms, which tend to be small, is associated to less employment in small registered plants, as one is the complementary of the other.

In order to measure the size of the informal economy, we compute the proportion of firms for which competition from the informal sector is a major or severe obstacle

[20]Transportation costs within a country may be better captured by the area of the country or a measure of internal distance, such as the one calculated by Mayer and Zignago (2011). It turns out that both measures are highly correlated with population in our sample of countries. Our results are robust to their inclusion.

to growth, as reported by ESWB. The advantage of this measure is that it can be calculated for each country-sector, as opposed to alternative measures -e.g. percentage of GDP-, such as, for instance, Djankov et al. (2002), which, moreover, are only available for a small sub-sample of countries.

We run the following regression:

$$\begin{aligned} SS_{ij} &= \beta_0 + \beta_1 \text{Doing Business}_i + \beta_2 \text{GDP pc}_i + \beta_3 \text{Informality}_{ij} \\ &\quad + \beta_4 \text{Log Population}_i + \sum_j \gamma_j + u_{ij} \end{aligned} \tag{5}$$

where SS_{ij} is the share of employment allocated to small plants in country i and sector j; Doing Business is the *Ease of Doing Business Index*; GDP pc is real per capita GDP with respect to the US; Informality is our measure of the informal economy; Log Population controls for total scale effects and γ_j are sector dummies corresponding to manufacturing, construction, services and trade; and u_{ij} is a disturbance term. We include sector dummies in order to control for technological characteristics of the sector that can influence the distribution of employment.

Table V reports the results of this regression. All columns represent regression results where the share of employment allocated to small plants is the dependent variable and standard errors are clustered at the country level. Columns (1), (2) and (3) represent regressions at the country-sector level as represented by equation 5. Columns (4) and (5) represent results for the regressions at the country level.[21]

In column (1) of Table V we observe that countries with a higher per capita GDP and a worse business environment are associated to lower shares of employment allocated to small plants, although the latter relationship is not statistically significant.

In column (2) we introduce the size of the informal economy and the total population of the country. The coefficient on the *Doing Business Index* increases by around 40 per cent, becoming statistically significant at 95 percent confidence level. The point estimate of .006 we find in the business environment means that in economies where distortions are very low (*Doing Business* = .01), we observe that the share of employment allocated to small plants is around 12 percentage points lower than that in a country where distortions are very high (*Doing Business* = 1.81).

[21] We report these results just for completeness.

The negative correlation between income per capita and the share of employment in small plants is .31. This means that countries with per capita GDP one standard deviation higher than the average are associated to a lower share of employment allocated to small plants of 4 percentage points. This is around 20 percent of the average of the dependent variable. We also find that larger countries tend to have less resources allocated to small plants. In particular, countries 1 percent bigger in terms of population are associated to a share of employment accounted by small plants 2 percentage points lower.

With respect to the size of the informal sector, we find that it is negatively correlated with the share of employment allocated to small registered plants. In particular, we find that in economies where competition of informal firms is very disruptive -proportion of firms for which competition of informal firms is an obstacle to growth = 1-, the share of employment in small registered plants is 12 percentage points lower than in those countries with no competition from informal firms -proportion of firms for which competition of informal firms is an obstacle to growth = 0-.

Finally, in column (3) we explore whether the association between the business environment and the size distribution of firms is more prevalent in different sectors. Hence, we interact the *Doing Business Index* with the sector dummies. We find that this association is specially acute in the retail & wholesale trade sector, followed by manufacturing and then by construction and services. Nevertheless, none of the interaction terms is significantly different from zero. However, the F tests of the *Doing Business* plus the corresponding interaction reveal that that the association between the business environment and the construction and services sectors is not statistically different from zero.

A Particular Type of Distortions: Financial Frictions

Financial frictions are a particular type of distortion that generates misallocation of resources. The mechanism through which financial frictions can generate misallocation is straightforward. Suppose that there are poor and rich entrepreneurs, and both rich and poor can be talented or untalented. In a context of lack of full enforcement of contracts, poor talented entrepreneurs will operate at a too small scale because they will not be able to capture enough resources from financial markets to achieve their optimal size. On the other hand, rich entrepreneurs will be able to finance themselves using

their own resources. Then, if the correlation between richness and talent is not one, misallocation arises in equilibrium: the marginal product of the factors of production used by poor talented entrepreneurs will be higher than the marginal product of factors used by untalented rich entrepreneurs. Hence, the economy would be better off in terms of output if factors are reallocated from the low marginal productivity entrepreneurs to the high marginal productivity ones.

Recent papers have emphasized that financial frictions can potentially generate high TFP losses through this mechanism. However, there is still some disagreement on the magnitude of these losses. Buera et al. (2011) calibrate a two sectors version of Lucas (1978) to the US economy, showing that financial frictions can generate TFP losses of up to 40 percent. Calibrating a similar model to micro-data of plants operating in South Corea, Midrigan and Xi (2010) find that financial frictions generate TFP losses of up to 7 percent. Both papers focus on steady state comparisons. Moll (2012) emphasizes that looking at transitions is crucial to measure the actual effect of financial frictions.

We use our cross-country size distribution of establishments to study the relationship between the share of employment allocated to small plants and financial frictions. We run regressions very similar to the ones presented above, but including only financial frictions as a measure of distortions.

Our measure of financial frictions relies on one component of the *Ease of Doing Business Index*, namely the *Getting Credit Index*. This index assesess the legal rights of lenders in performing financial transactions and the availability of credit histories information. To do so, it uses two set of indicators. The first one evaluates how well collateral and bankruptcy laws facilitates lending. The second measures the coverage, scope and accessibility of credit information available through public credit registries and private credit bureaus.[22] The *Getting Credit Index* ranks countries from lower to higher financial frictions, hence a higher amount of the index means a higher level of financial constraints.

Table VIII shows the association between the share of labor employed in small plants and the *Getting Credit Index*, controlling for the same covariates as in Equation (5). We observe a robust negative correlation between financial constraints and the share of labor allocated to small plants, significant at the 99 percent confidence level. Moreover, the magnitude of this correlation is bigger than the one associated to the *Doing*

[22]See http://www.doingbusiness.org/methodology/getting-credit#legalRights.

Business Index. For instance, the coefficient in column (2) means that, conditional on covariates, a country with very high financial frictions -*Getting Credit* = 1.81- exhibits 15 percentage points less share of labor in small plants than a country with very low financial frictions -*Getting Credit* = .01-.

The coefficients associated to the rest of the covariates have the same sign and similar magnitude as the ones found in Table V. Finally, the interaction terms show that the association of financial constraints and the size distribution is strongest in the trade sector, which is significantly different from that of manufacturing. This suggests that the presence of financial frictions distorts more, relative to other sectors, the size distribution of plants in wholesale and retailing.

6 The Missing Middle

As mentioned in the Introduction, some researchers have observed an interesting cross-country pattern in the manufacturing size distribution of plants. Low-income countries tend to allocate almost all the employment to either small or large plants. This contrasts with high income countries, where the share of employment accounted by plants is monotonically increasing in size. This result is presented in Table 1 in Tybout (2000).

Some works have provided different theories in order to rationalize this phenomenon. In an old work, Rauch (1991) constructs a Lucas (1978) type of model that generates a missing middle in the size distribution of plants. In his model, entrepreneurs choose which sector to operate in, either informal or formal. In the informal sector they avoid regulation.[23] However, in order to keep themselves out of the hand of the government, they must remain below a given size threshold. Thus, those entrepreneurs that decide to be formal always operate very above the threshold because it never pays to be just large enough to suffer regulation. This is also the case in Gallipoli and Goyette (2011), where a significant mass of small firms remain just below the threshold at which the audit intensity sharply increases. These papers show that standard models of heterogenous plants are able to reproduce the missing middle with the introduction of distortions.

In this section, we add new empirical evidence of this phenomenon. In particular, we want to provide descriptive evidence about how much the distribution of employment across firms in low-income countries deviate from the one in high income countries. To this end, we run two different exercises. First, exploiting two very broad datasets, we try to compare in very detail the size distribution of Germany and India. These datasets contains a very large number of observations which allows us to run a very precise comparison, looking carefully at the whole distribution. Second, in order to perform a broader cross-country comparison, we compare the size distribution of countries calculated from the ESWB to that reported by Tybout (2000).

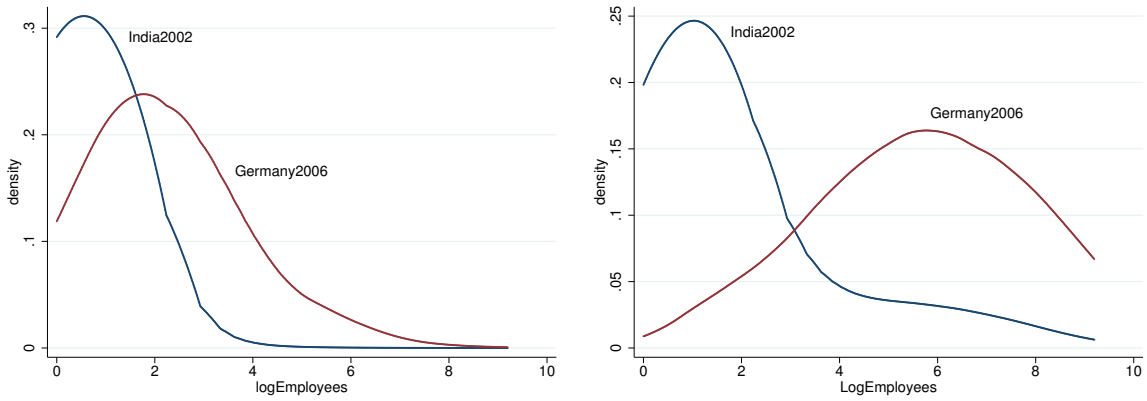
Comparing two different economies: India vs Germany

Although the 'missing middle' is a highly discussed phenomenon in Development Economics, it is not straightforward to identify it in the data. The main reason is that

[23] Rauch (1991) assumes the existence of a minimum wage in the formal sector.

most of datasets at the plant/firm level do not include small production units in their sample frames, which is translated into an overestimation of the importance of middle firms. In order to control for this problem, we use a dataset of the Indian manufacturing sector that allows us to measure small production units: the National Sample Survey (NSS). This data set is conducted every five years as one of the modules of the Indian National Sample Survey. In particular, we use the 56^{th} round of NSS, whose sampling period is 2000-01. We also use the widely used Annual Survey of Industries (ASI) for the same year, which covers larger plants. Both datasets are complementary to each other, which means that combining them allows us to have a representative sample of the whole distribution of plants in the Indian manufacturing sector.[24]. Then, we take Germany as an undistorted benchmark economy and compare its manufacturing size distribution to the Indian one.[25]

FIGURE IV
GERMANY VS INDIA



(a) Distribution of production units

(b) Distribution of employment

In figure VI we plot both the distribution of production units and employment for India and Germany. As expected, the Indian distribution of production units is much

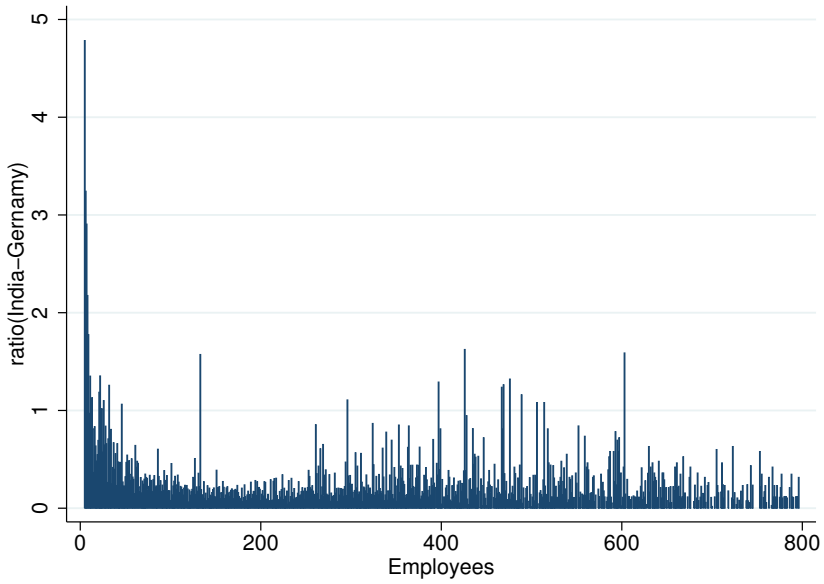
[24]The 56^{th} NSS contains around 150,000 observations. The ASI contains around 50,000 for the year we use. Of course, all observations are appropriately weighted, which allows to estimate population statistics

[25]For Germany we use AMADEUS dataset for the year 2006, which contains around 105,000 observations for the manufacturing sector. AMADEUS contains representative samples of non-government owned European firms, including very small ones

more skewed towards to the right than the German one. This pattern becomes dramatic in the distribution of employment: the distribution of employment in Germany is a bit skewed towards the right whereas in India it is skewed towards the left. This finding is consistent with those found by Poschke (2011) and also with our findings in section 4. However, we cannot say much about the missing middle from these figures. Then, in order to identify the missing middle, we plot the ratio of employment shares between the two countries. To avoid the rather arbitrary decision on the construction of size bins, we compute the ratio for each possible firm. From figure V we see that the share of employment accounted by plants of small size is much higher in India than in Germany. However, we observe that as size starts to increase, the ratio starts to be lower and then it increases again (around 300 employees). This is evidence of big plants in India accounting for a share of employment as high as big plants in Germany. So, in other words, there is too much employment allocated to small plants and too little allocated to medium plants in India, relative to Germany. From figure VI we can clearly see this bimodality in the size distribution of employment.[26]

Figure V

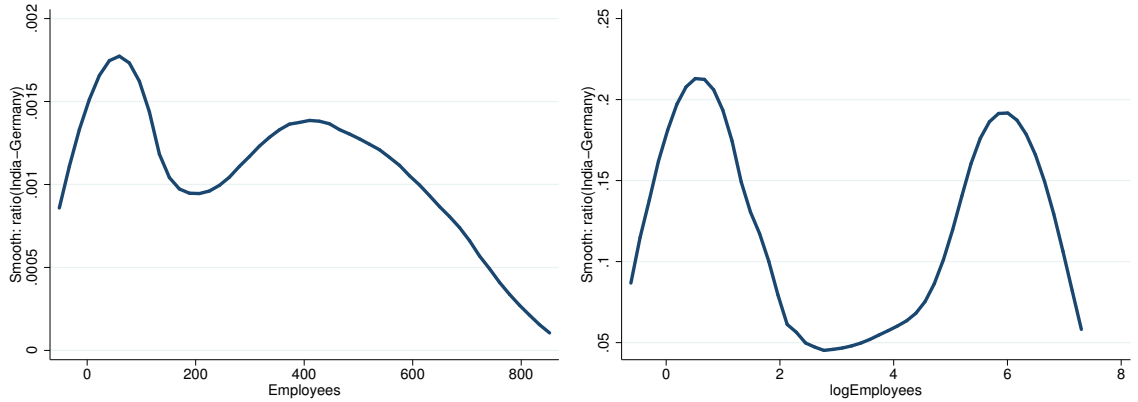
Comparison in the Size Distribution of Employment: India vs Germany



[26]Figures (a) and (b) in VI are smooth versions of figure V where the latter is plotted against the log scale.

Figure VI

Comparison in the Size Distribution of Employment: India vs Germany



(a) (b)

The Missing Middle in the ESWB countries

Tybout (2000) is the first paper stressing the fact that in developing countries labor allocated to firms of medium size is lower relative to thal allocated to small and large firms. This phenomenon is mostly documented from the estimations performed in Liedholm and Mead (1987) corresponding mostly for a bunch of developing countries in the 1970s. This paper calculates the share of labor allocated to manufacturing plants of less than 9 employees, between 10 and 49 employees, and more than 49 employees. As already mentioned, the fact that firms between 10 and 49 employees account for a remarkably low share of employment was underlined by Tybout (2000) giving birth to the so-called 'missing middle' phenomenon.

In Table VI we redo the calculations performed by Liedholm and Mead (1987) using our database, focusing on the countries they considered. The first line of each country corresponds to the numbers computed by Liedholm and Mead (1987) and showed in Table 1 of Tybout (2000). We also consider additional sectors beyond manufacturing, such as services and trade.

As it is patent, our estimations of the labor accounted by medium firms are much larger than those found by Liedholm and Mead (1987). Therefore, the missing mid-

dle phenomenon is clearly not prevalent in our database. Although around 35 years separate the distributions calculated by Liedholm and Mead (1987) and our calculations, and these could explain part of the difference, this is more likely to be explained by the characteristics of the database. As already mentioned, ESWB do not target those establishment with less than 5 people employed, which are included in Liedholm and Mead (1987). Thus, as stressed in the Introuduction, our statistics of the size distribution of establishments are biased towards less labor accounted by small firms, and therefore, more labor accounted by medium and large plants. Then, the missing middle phenomenon is much harder to emerge. On the contrary, the surveys used in Liedholm and Mead (1987) do not truncate the size distribution, making them more suitable to analyse the existence of the missing middle. This exercise then stresseses the fact that left-truncated survey or census data, common in the literature of missalocation of resources, is less appropiate to analyse phenomena *within* countries, rather than performing a cross-country comparison.

TABLE VI

DISTRIBUTION OF EMPLOYMENT ACROSS PLANT SIZES: THE MISSING MIDDLE

Country	Year	Sector	Number of Workers		
			1 - 9	10 - 49	> 49
Tanzania	1967	Manufacturing	0.56	0.07	0.37
	2006	Manufacturing	0.03	0.17	0.80
	2006	Services	0.11	0.41	0.48
	2006	Trade	0.27	0.63	0.10
Ghana	1970	Manufacturing	0.84	0.01	0.15
	2007	Manufacturing	0.05	0.20	0.76
	2007	Services	0.21	0.53	0.26
	2007	Trade	0.32	0.49	0.19
Kenya	1969	Manufacturing	0.49	0.10	0.41
	2007	Manufacturing	0.00	0.05	0.95
	2007	Services	0.05	0.23	0.72
	2007	Trade	0.08	0.29	0.62
Sierra Leone	1974	Manufacturing	0.90	0.05	0.05
	2009	Manufacturing	0.34	0.54	0.12
	2009	Services	0.17	0.20	0.63
	2009	Trade	0.32	0.28	0.39
Indonesia	1977	Manufacturing	0.77	0.07	0.16
	2009	Manufacturing	0.14	0.16	0.70
	2009	Services	0.08	0.27	0.64

Table VI

Distribution of Employment Across Plant Sizes: The Missing Middle (Continued)

Country	Year	Sector	Number of Workers		
			1 - 9	10 - 49	> 49
	2009	Trade	0.18	0.36	0.46
Zambia	1985	Manufacturing	0.83	0.01	0.16
	2007	Manufacturing	0.01	0.09	0.91
	2007	Services	0.03	0.20	0.77
	2007	Trade	0.13	0.24	0.63
Honduras	1979	Manufacturing	0.68	0.08	0.24
	2010	Manufacturing	0.06	0.17	0.77
	2010	Services	0.04	0.10	0.85
	2010	Trade	0.05	0.13	0.82
Philippines	1974	Manufacturing	0.66	0.05	0.29
	2009	Manufacturing	0.01	0.10	0.89
	2009	Services	0.01	0.19	0.80
	2009	Trade	0.02	0.09	0.89
Nigeria	1972	Manufacturing	0.59	0.26	0.15
	2007	Manufacturing	0.07	0.38	0.55
	2007	Services	0.16	0.57	0.27
	2007	Trade	0.32	0.57	0.10
Jamaica	1978	Manufacturing	0.35	0.16	0.49
	2010	Manufacturing	0.01	0.26	0.72
	2010	Services	0.05	0.25	0.70
	2010	Trade	0.03	0.53	0.44
Colombia	1973	Manufacturing	0.52	0.13	0.35
	2006	Manufacturing	0.04	0.36	0.60
	2006	Services	0.09	0.49	0.42
	2006	Trade	0.08	0.31	0.60

Table VI shows the distribution of employment for a selected sample of countries in the manufacturing, services and trade sectors. Data before 2006 come from Liedholm and Mead (1987). See the Appendix and the text for further details.

7 Conclusions

Government policies distort the way resources are allocated across heterogeneous production units. These distortions can be a source of aggregate productivity differences across countries. A recent strand of literature identifies specific policies that prevent firms from growing, generating high efficiency losses. We add to this literature by performing a cross-country analysis. We use comparable firm-level surveys of 104 countries to study how the size distribution of plants varies across economies in different stages of development.

We find that small plants tend to be less productive than large establishments, suggesting that a higher amount of labor allocated to small establishments is indicative of a higher degree of misallocation of resources. Differences in income per capita accounts for a portion of the variation in size distribution, however they do not provide the whole story. On the contrary, we provide evidence that a better business environment is associated to a lower share of labor allocated to small plants, even after controlling for income per capita. Of the business regulations associated to differences in the size distribution, we find that those facilitating the functioning of financial markets are specially important.

Our database allows us also to disentangle how this pattern varies across sectors. We find that misallocation is decreasing with income per capita not only in manufacturing, but also in trade and services. Interestingly, the association between financial frictions and size distribution is magnified in the retail and wholesale sector. This heterogeneous effect of distortions across sectors is an interesting result. We leave a deeper exploration of this issue for further research.

References

Alfaro, Laura, Andrew Charlton, and Fabio Kanczuk, "Plant-Size Distribution and Cross-Country Income Differences," June 2008. NBER Working Paper 14060.

Banarji, R, "Average Size of Plants in Manufacturing and Capital Intensity. A Cross-country Analysis by Industry,Ó *Journal of Development Economics*, 1978, *9* (5),155-166.

Bollard, Albert, Peter Klenow, and Gunjan Sharma, "India's Mysterious Manufacturing Miracle," July 2011. Working Paper.

Buera, Francisco, Joseph Kaboski, and Yongseok Shin, "Finance and Development: A Tail of Two Sectors," *American Economic Review*, 2011, *101.*

_ **and Alexander Monge-Naranjo**, "Foreign Know-How, Firm Control and the Income of Developing Countries," *Quaterly Journal of Economics*, 2009, *124.*

Caselli, F, "Accounting for Income Differences across Countries,Ó in *Handbook of Economic Growth* Vol. 1 A, P. Aghion and S. Durlauf, eds. (Amsterdam: Elsevier, 2005, Chap.9)

Dasgupta, K, "The Missing Middle in Developing Countries RevisitedÓ 2010, Working Paper

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, "The Regulation of Entry," *Quarterly Journal of Economics*, 2002, *117.*

_ **and Ana Hidalgo-Cabrillana**, "On Finance as a Theory of TFP, Cross-Industry Productivity Differences, and Economic Rents", *International Economics Review*, 2008, *49* (2), 437–473

_ **Gallipoli and J Goyette**, "Distortions, Efficiency and the Size Disribution of Firms," 2011. Working paper.

Garcia-Santana, Manuel, "Foreign Ownership, Financial Frictions and Size Distribution of Plants," November 2011. Working Paper.

_ **and Josep Pijoan-Mas**, "The Small Scale Reservation Law and the Misallocation of Talent," February 2011. CEPR Discussion Paper, 8242.

Guner, Nezih, Gustavo Ventura, and Xu Yi, "Macroeconomic Implications of Size-Dependent Policies," February 2007. CEPR Discussion Paper no. 6138.

Guner, Nezih, Gustavo Ventura, and Xu Yi, "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, October 2008, *11* (4), 721–744.

_ **and Charles Jones**, "Why Do Some Countries Produce So Much More Output per Worker Than Others?", *Quaterly Journal of Economics*, 1999, *114*.

Heston, Alan, Robert Summers, and Bettina Aten, "Penn World Table Version 7.0," May 2011. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Klenow, Peter and Andres. Rodriguez-Clare, "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in *NBER Macroeconomics Annual*, B. Bernanke and J. Rotemberg, eds. (Cambridge, MA: MIT Press, 1997)

Mayer, Thierry and Soledad. Zignago, "Notes on CEPII's Distances Measures: the GeoDist Database" ,2011, Working paper

Melitz, M,"The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity" *Econometrica*, 2003, *71* , 1695-1725.

Midrigan, Virgiliu and Daniel Xu, "Finance and Misallocation: Evidence from Plant Level Data," 2010, Working Paper

Mohan, R, "Small-Scale Industry Policy In India: a Critical Evaluation,Ó in *in Economic Policy Reforms and the Indian Economy* ed. by A. Krueger, chap. 6. University of Chicago Press.

Moll, B, "Productivity Losses from Capital Frictions: Can Self-Financing Undo Capital Misallocation?,Ó 2012, Working Paper

Lewis, W, "The Power of Productivity: Wealth, Poverty, and the Threat to Global Stability,Ó University of Chicago Press.

Liedholm, Carl and Donald. Mead, "Small Scale Industries in Developing Countries: : Empirical Evidence and Policy Implications," *http://www.aec.msu.edu/fs2/papers/older/idp9.pdf*, Mimeo

Lucas, R, "On the Size Distribution of Business Firms,Ó *Bell Journal of Economics*, 1978, *9* (2), 508-523.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacuring TFP in China and India," *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Poschke, M, "The Firm Size Distribution across Countries and Skill-biased Change in Entrepreneurial TechnologyÓ 2012, Working Paper

Rauch, J, "Modeling the Informal Sector Formally,Ó *Journal of Development Economics*, 1991, *35* , 33-47.

Restuccia, Diego and Richard Rogerson, "Policy Distorsions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, October 2008, *11* (4), 707–720.

de Soto, Hernando, "The Other Path: The In- visible Revolution in the Third World," *NY: Har- per-Row.*

Tybout, James R., "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?," *Journal of Economic Literature*, March 2000, *38*, 11–44.

Table IV

Share of Employment Accounted by Small, Medium and Large Firms Across Countries: All Firms

	Small Firms			Medium Firms			Large Firms		
	Low Income	Low Middle Income	Upper Middle & High Inc.	Low Income	Low Middle Income	Upper Middle & High Income	Low Income	Low Middle Income	Upper Middle & High Income
Manufacturing	0.18	0.13	0.09	0.29	0.28	0.26	0.54	0.59	0.65
	(0.17)	(0.12)	(0.06)	(0.14)	(0.16)	(0.10)	(0.26)	(0.24)	(0.15)
Construction	0.19	0.14	0.08	0.34	0.35	0.30	0.47	0.51	0.62
	(0.26)	(0.22)	(0.08)	(0.25)	(0.24)	(0.22)	(0.35)	(0.29)	(0.25)
Services	0.30	0.29	0.17	0.36	0.38	0.29	0.34	0.33	0.53
	(0.23)	(0.25)	(0.13)	(0.20)	(0.24)	(0.15)	(0.25)	(0.28)	(0.23)
Trade	0.43	0.30	0.20	0.31	0.32	0.27	0.27	0.38	0.53
	(0.23)	(0.22)	(0.12)	(0.15)	(0.16)	(0.13)	(0.26)	(0.26)	(0.23)

Table IV shows the average share of employment accounted by small, medium and large establishments across countries according to their level of income. Standard deviations in parenthesis. See the Appendix and the text for further details.

Table V

Share of Employment Accounted by Small Firms and Ease of Doing Business Index

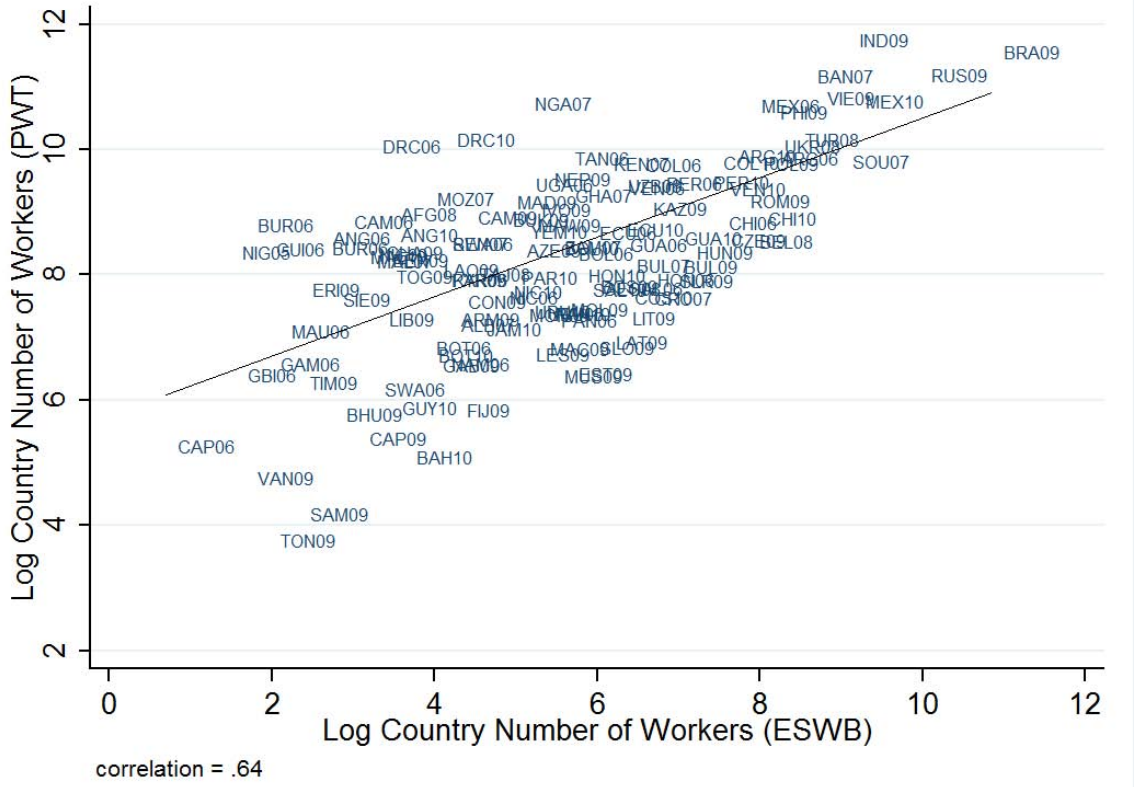
	(1)	(2)	(3)	(4)	(5)
Dep. Variable: Share of Employment in Small Firms					
Ease of Doing Business Index	0.0438	0.0595**	0.0539**	0.0470*	0.0663***
	(0.0298)	(0.0284)	(0.0242)	(0.0272)	(0.0251)
Doing Business * Construction			-0.0184		
			(0.0347)		
Doing Business * Services			-0.0136		
			(0.0346)		
Doing Business * Trade			0.0532		
			(0.0344)		
GDP per capita	-0.2952***	-0.3063***	-0.3066***	-0.1637*	-0.1742**
	(0.0969)	(0.0905)	(0.0907)	(0.0838)	(0.0769)
Informal Economy		-0.1200*	-0.1228*		-0.1227*
		(0.0693)	(0.0700)		(0.0649)
Log Population		-0.0206**	-0.0205**		-0.0247***
		(0.0086)	(0.0087)		(0.0089)
Constant				0.1346***	0.3763***
				(0.0418)	(0.0972)
Sector Dummies	YES	YES	YES	NO	NO
F Test Construction			[0.3874]		
F Test Services			[0.3003]		
F Test Trade			[0.0093]		
Observations	490	490	490	125	125
R-squared	0.20	0.25	0.25	0.10	0.24

Table V shows the regressions of the share of employment accounted by small firms against the *Ease of Doing Business Index*. Column (1) to (3) are regressions at the country-sector level whereas columns (4) and (5) are at the country level. F Test corresponds to the p-value of a test of the coefficient of *Ease of Doing Business Index* plus the corresponding interaction not being signficantly different from zero. Standard errors are in parenthesis, clustered at the country-survey level in columns (1) to (3) and at the country level in columns (4) and (5). Significance levels: *: 10%; **: 5%; ***: 1%. See Appendix A and the text for further details.

Figure VII

Number of Workers Implied by ESWB Compared to Penn World Tables

Panel A: Raw Correlation



Panel B: Partial Correlation

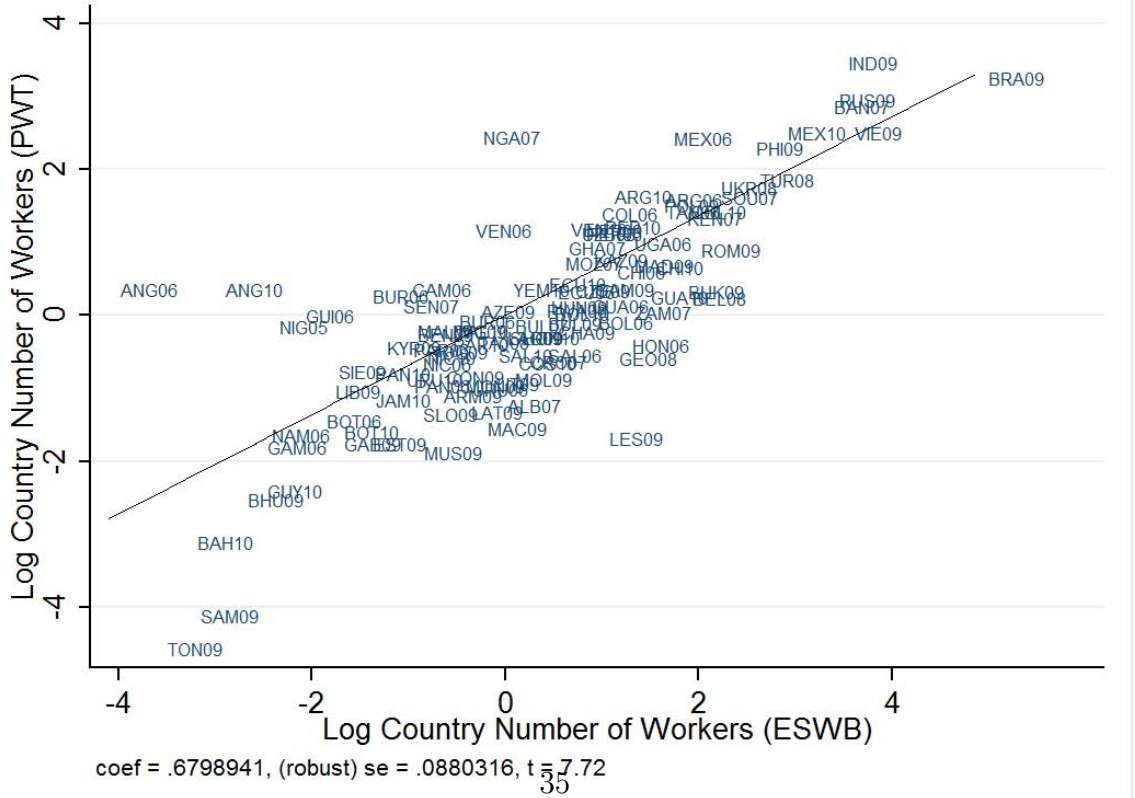


Table VII

Share of Employment Accounted by Small, Medium and Large Firms Across Countries: Domestic Firms

	Small Firms			Medium Firms			Large Firms		
	Low Income	Low Middle Income	Upper Middle & High Inc.	Low Income	Low Middle Income	Upper Middle & High Income	Low Income	Low Middle Income	Upper Middle & High Income
Manufacturing	0.23	0.17	0.11	0.35	0.31	0.29	0.43	0.52	0.60
	(0.19)	(0.13)	(0.08)	(0.13)	(0.16)	(0.10)	(0.25)	(0.25)	(0.17)
Construction	0.24	0.15	0.09	0.36	0.36	0.33	0.40	0.49	0.59
	(0.29)	(0.22)	(0.09)	(0.25)	(0.24)	(0.23)	(0.34)	(0.29)	(0.27)
Services	0.34	0.32	0.21	0.36	0.41	0.33	0.30	0.27	0.47
	(0.25)	(0.26)	(0.15)	(0.19)	(0.25)	(0.19)	(0.24)	(0.28)	(0.27)
Trade	0.48	0.32	0.22	0.31	0.35	0.30	0.22	0.33	0.47
	(0.22)	(0.22)	(0.13)	(0.15)	(0.18)	(0.14)	(0.22)	(0.27)	(0.24)

Table VII shows the average share of employment allocated to domestic firms accounted by small, medium and large firms across countries according to the level of income. Standard deviations in parenthesis. See Appendix and the text for further details.

Table VIII

Share of Employment Accounted by Small Firms and Getting Credit Index

	(1)	(2)	(3)	(4)	(5)
Dep. Variable: Share of Employment in Small Firms					
Getting Credit Index	0.0852***	0.0839***	0.0557***	0.0710***	0.0677***
	(0.0218)	(0.0218)	(0.0184)	(0.0195)	(0.0190)
Getting Credit * Construction			0.0348		
			(0.0353)		
Getting Credit * Services			-0.0107		
			(0.0350)		
*Getting Credit*s * Trade			0.0866**		
			(0.0343)		
GDP per capita	-0.2219***	-0.2587***	-0.2591***	-0.1270*	-0.1721**
	(0.0726)	(0.0714)	(0.0717)	(0.0690)	(0.0718)
Informal Economy		-0.1299*	-0.1313*		-0.1210*
		(0.0679)	(0.0681)		(0.0629)
Log Population		-0.0190*	-0.0189*		-0.0237**
		(0.0096)	(0.0097)		(0.0097)
Constant				0.1202***	0.3833***
				(0.0228)	(0.0928)
Sector Dummies	YES	YES	YES	NO	NO
F Test Construction			[0.0091]		
F Test Services			[0.2107]		
F Test Trade			[0.0001]		
Observations	462	462	462	118	118
R-squared	0.24	0.28	0.29	0.15	0.27

Table VIII shows the regressions of the share of employment accounted by small firms against the *Getting Credit Index*. Column (1) to (3) are regressions at the country-sector level whereas columns (4) and (5) are at the country level. F Test corresponds to the p-value of a test of the coefficient of *Getting Credit Index* plus the corresponding interaction not being signficantly different from zero. Standard errors are in parenthesis, clustered at the country-survey level in columns (1) to (3) and country level in columns (3) to (5). Significance levels: *: 10%; **: 5%; ***: 1%. See Appendix A and the text for further details.

Appendix

A Data: Definitions and Sources

TABLE A.1
DATA DEFINITIONS AND SOURCES

VARIABLE	SOURCE	DEFINITION
Size of an Establishment (Employees)	ESWB	Permanent (l1) + Temporary (l6) Employees in last fiscal year
Manufacturing	ESWB	ISIC (Rev. 3.1) between 15 and 37
Construction	ESWB	ISIC (Rev. 3.1) 45
Services	ESWB	ISIC (Rev. 3.1) 40, 41 and between 55 and 95
Trade	ESWB	ISIC (Rev. 3.1) 50, 51 and 52
Number of Workers	PWT 7.0	$\frac{\text{PPP Converted GDP Per Capita}}{\text{PPP Converted GDP Chain per worker}}$Population
Employment in Agriculture	World Bank	Share of employment in agriculture as percentage of total employment
Small Establishment	ESWB	Size of less than 20 employees
Medium Establishment	ESWB	Size between 20 and 99 employees
Large Establishment	ESWB	Size of 100 or more employees
Foreign Ownership	ESWB	50 per cent or more of the firm is owned by private foreign individuals, companies or organizations (b2b)
Ease of Doing Business Index	World Bank	Ranking of business environment.
GDP per Capita	World Bank	GDP per Capita, PPP (constant 2005 international dollars) relative to the US
Getting Credit Index	World Bank	Ranking of financial constraints.
Informal Economy	ESWB	Percentage of firms for which competition of informal firms is a major or severe obstacle to growht (e30)
Log Population	PWT 7.0	Log Population
Productivity	ESWB	$\frac{\text{Sales (d2)}-\text{Cost of Raw Materials (n2e)}-\text{Cost of Electricity (n2b)}}{\text{Number of Employees}}$
Export Status	ESWB	Establishment sales part of their output abroad (d3b,d3c)
Age	ESWB	Age of the Establishment (b5)

Table A.1 shows definitions and sources of all variables used throughout the analysis. When the Source is the Enterprise Surveys of the World Bank (ESWB) the codes in parenthesis in the Definition correspond to the code of the Questionnarie of the ESWB.

B Share of Employment Accounted by Small, Medium and Large Firms: Full Sample of Countries. Manufacturing

TABLE B.1
SHARE OF EMPLOYMENT ACCOUNTED BY SMALL, MEDIUM AND LARGE FIRMS: MANUFACTURING

Country	Year	Small Firms	Medium Firms	Large Firms
Low Income Countries				
Afghanistan	2008	0.10	0.39	0.51
Bangladesh	2007	0.01	0.06	0.93
Benin	2009	0.12	0.14	0.73
BurkinaFaso	2006	0.08	0.22	0.70
BurkinaFaso	2009	0.12	0.24	0.64
Burundi	2006	0.19	0.30	0.51
Chad	2009	0.13	0.25	0.62
DRC	2006	0.15	0.34	0.51
DRC	2010	0.12	0.25	0.63
Eritrea	2009	0.26	0.51	0.23
Gambia	2006	0.11	0.53	0.37
Guinea	2006	0.29	0.19	0.52
GuineaBissau	2006	0.52	0.48	0.00
Kenya	2007	0.01	0.11	0.88
Kyrgyz Republic	2009	0.05	0.25	0.70
Liberia	2009	0.46	0.54	0.00
Madagascar	2009	0.03	0.14	0.84
Malawi	2009	0.02	0.18	0.80
Mali	2007	0.26	0.40	0.34
Mali	2010	0.21	0.33	0.46
Mozambique	2007	0.21	0.47	0.32
Nepal	2009	0.33	0.36	0.31
Niger	2005	0.07	0.21	0.72
Niger	2009	0.37	0.39	0.24
Rwanda	2006	0.01	0.08	0.91
Sierra Leone	2009	0.73	0.20	0.07
Tajikistan	2008	0.03	0.19	0.77
Tanzania	2006	0.08	0.24	0.68
Togo	2009	0.15	0.40	0.45
Uganda	2006	0.08	0.20	0.72
Average		0.18	0.29	0.54
Standard Dev.		(0.17)	(0.14)	(0.26)
Low Middle Income Countries				

Table B.1

Share of Employment Accounted by Small, Medium and Large Firms: Manufacturing (Continued)

Country	Year	Small Firms	Medium Firms	Large Firms
Angola	2006	0.52	0.29	0.20
Angola	2010	0.12	0.26	0.62
Armenia	2009	0.05	0.26	0.69
Bhutan	2009	0.13	0.29	0.59
Bolivia	2006	0.09	0.37	0.54
Bolivia	2010	0.05	0.41	0.53
Cameroon	2006	0.02	0.14	0.84
Cameroon	2009	0.05	0.15	0.80
CapeVerde	2006	0.25	0.65	0.10
CapeVerde	2009	0.27	0.53	0.20
Congo	2009	0.13	0.49	0.37
ElSalvador	2006	0.07	0.13	0.80
ElSalvador	2010	0.05	0.16	0.79
Fiji	2009	0.06	0.24	0.70
Georgia	2008	0.06	0.13	0.81
Ghana	2007	0.11	0.17	0.72
Guatemala	2006	0.21	0.20	0.59
Guatemala	2010	0.05	0.14	0.81
Guyana	2010	0.02	0.18	0.81
Honduras	2006	0.11	0.18	0.71
Honduras	2010	0.10	0.17	0.73
Indonesia	2009	0.22	0.15	0.63
Ivory Coast	2009	0.14	0.25	0.61
Kosovo	2009	0.21	0.38	0.41
LaoPDR	2009	0.15	0.22	0.63
Lesotho	2009	0.01	0.02	0.98
Mauritania	2006	0.17	0.45	0.38
Micronesia	2009	0.32	0.68	0.00
Moldova	2009	0.07	0.22	0.71
Mongolia	2009	0.07	0.40	0.54
Nicaragua	2006	0.22	0.29	0.49
Nicaragua	2010	0.07	0.13	0.80
Nigeria	2007	0.20	0.47	0.32
Paraguay	2006	0.10	0.36	0.54
Paraguay	2010	0.06	0.27	0.67
Philippines	2009	0.04	0.17	0.79
Samoa	2009	0.09	0.25	0.66
Senegal	2007	0.10	0.14	0.76
Swaziland	2006	0.02	0.06	0.93
Timor Leste	2009	0.29	0.50	0.21
Tonga	2009	0.50	0.50	0.00
Ukraine	2008	0.04	0.16	0.80

Table B.1

Share of Employment Accounted by Small, Medium and Large Firms: Manufacturing (Continued)

Country	Year	Small Firms	Medium Firms	Large Firms
Uzbekistan	2008	0.17	0.35	0.48
Vanuatu	2009	0.21	0.54	0.25
Vietnam	2009	0.01	0.11	0.88
Yemen	2010	0.28	0.17	0.55
Zambia	2007	0.02	0.18	0.80
Average		0.13	0.28	0.59
Standard Dev.		(0.12)	(0.16)	(0.24)
Upper Middle & High Income Countries				
Albania	2007	0.11	0.43	0.46
Argentina	2006	0.04	0.14	0.82
Argentina	2010	0.05	0.21	0.74
Azerbaijan	2009	0.08	0.20	0.72
Bahamas	2010	0.16	0.47	0.38
Belarus	2008	0.01	0.08	0.91
Bosnia and Herzegovina	2009	0.06	0.35	0.59
Botswana	2006	0.08	0.25	0.67
Botswana	2010	0.08	0.29	0.63
Brazil	2009	0.02	0.19	0.79
Bulgaria	2007	0.08	0.30	0.62
Bulgaria	2009	0.13	0.25	0.62
Chile	2006	0.03	0.17	0.80
Chile	2010	0.03	0.22	0.75
Colombia	2006	0.13	0.53	0.34
Colombia	2010	0.05	0.14	0.81
Costarica	2010	0.06	0.21	0.73
Croatia	2007	0.09	0.26	0.65
Czech Republic	2009	0.04	0.26	0.70
Ecuador	2006	0.09	0.26	0.65
Ecuador	2010	0.06	0.15	0.79
Estonia	2009	0.14	0.39	0.47
Fyr Macedonia	2009	0.09	0.33	0.59
Gabon	2009	0.19	0.31	0.50
Grenada	2010	0.23	0.53	0.24
Hungary	2009	0.05	0.18	0.77
Jamaica	2010	0.11	0.35	0.54
Kazakhstan	2009	0.04	0.20	0.77
Latvia	2009	0.10	0.41	0.48
Lithuania	2009	0.12	0.35	0.53
Mauritius	2009	0.08	0.24	0.68
Mexico	2006	0.09	0.18	0.73
Mexico	2010	0.07	0.15	0.78

Table B.1

Share of Employment Accounted by Small, Medium and Large Firms: Manufacturing (Continued)

Country	Year	Small Firms	Medium Firms	Large Firms
Montenegro	2009	0.26	0.37	0.37
Namibia	2006	0.07	0.20	0.73
Panama	2006	0.10	0.29	0.61
Panama	2010	0.11	0.33	0.56
Peru	2006	0.03	0.12	0.84
Peru	2010	0.04	0.15	0.80
Poland	2009	0.06	0.15	0.79
Romania	2009	0.08	0.22	0.70
Russia	2009	0.01	0.12	0.88
Serbia	2009	0.05	0.27	0.68
Slovak Republic	2009	0.13	0.24	0.63
Slovenia	2009	0.07	0.20	0.74
SouthAfrica	2007	0.05	0.27	0.68
Turkey	2008	0.07	0.21	0.72
Uruguay	2006	0.26	0.43	0.31
Uruguay	2010	0.15	0.31	0.54
Venezuela	2006	0.21	0.32	0.47
Venezuela	2010	0.09	0.22	0.68
Average		0.09	0.26	0.65
Standard Dev.		(0.06)	(0.10)	(0.15)

Table B.1 shows the average share of employment in the manufacturing sector accounted by small, medium and large firms across countries included in the Standardized Data 2006-2010 of the ESWB. See the Appendix and the text for further details.

CEMFI WORKING PAPERS

0801 *David Martinez-Miera and Rafael Repullo*: "Does competition reduce the risk of bank failure?".

0802 *Joan Llull*: "The impact of immigration on productivity".

0803 *Cristina López-Mayán:* "Microeconometric analysis of residential water demand".

0804 *Javier Mencía and Enrique Sentana:* "Distributional tests in multivariate dynamic models with Normal and Student *t* innovations".

0805 *Javier Mencía and Enrique Sentana:* "Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation".

0806 *Dante Amengual and Enrique Sentana:* "A comparison of mean-variance efficiency tests".

0807 *Enrique Sentana:* "The econometrics of mean-variance efficiency tests: A survey".

0808 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla:* "Are joint negotiations in standard setting "reasonably necessary"?".

0809 *Rafael Repullo and Javier Suarez:* "The procyclical effects of Basel II".

0810 *Ildefonso Mendez:* "Promoting permanent employment: Lessons from Spain".

0811 *Ildefonso Mendez:* "Intergenerational time transfers and internal migration: Accounting for low spatial mobility in Southern Europe".

0812 *Francisco Maeso and Ildefonso Mendez:* "The role of partnership status and expectations on the emancipation behaviour of Spanish graduates".

0813 *Rubén Hernández-Murillo, Gerard Llobet and Roberto Fuentes:* "Strategic online-banking adoption".

0901 *Max Bruche and Javier Suarez:* "The macroeconomics of money market freezes".

0902 *Max Bruche:* "Bankruptcy codes, liquidation timing, and debt valuation".

0903 *Rafael Repullo, Jesús Saurina and Carlos Trucharte*: "Mitigating the procyclicality of Basel II".

0904 *Manuel Arellano and Stéphane Bonhomme*: "Identifying distributional characteristics in random coefficients panel data models".

0905 *Manuel Arellano, Lars Peter Hansen and Enrique Sentana*: "Underidentification?".

0906 *Stéphane Bonhomme and Ulrich Sauder*: "Accounting for unobservables in comparing selective and comprehensive schooling".

0907 *Roberto Serrano*: "On Watson's non-forcing contracts and renegotiation".

0908 *Roberto Serrano and Rajiv Vohra*: "Multiplicity of mixed equilibria in mechanisms: a unified approach to exact and approximate implementation".

0909 *Roland Pongou and Roberto Serrano*: "A dynamic theory of fidelity networks with an application to the spread of HIV / AIDS".

0910 *Josep Pijoan-Mas and Virginia Sánchez-Marcos:* "Spain is different: Falling trends of inequality".

0911 *Yusuke Kamishiro and Roberto Serrano:* "Equilibrium blocking in large quasilinear economies".

0912 *Gabriele Fiorentini and Enrique Sentana:* "Dynamic specification tests for static factor models".

0913 *Javier Mencía and Enrique Sentana:* “Valuation of VIX derivatives”.

1001 *Gerard Llobet and Javier Suarez:* “Entrepreneurial innovation, patent protection and industry dynamics”.

1002 *Anne Layne-Farrar, Gerard Llobet and A. Jorge Padilla:* “An economic take on patent licensing: Understanding the implications of the “first sale patent exhaustion” doctrine.

1003 *Max Bruche and Gerard Llobet:* “Walking wounded or living dead? Making banks foreclose bad loans”.

1004 *Francisco Peñaranda and Enrique Sentana:* “A Unifying approach to the empirical evaluation of asset pricing models”.

1005 *Javier Suarez:* “The Spanish crisis: Background and policy challenges”.

1006 *Enrique Moral-Benito*: “Panel growth regressions with general predetermined variables: Likelihood-based estimation and Bayesian averaging”.

1007 *Laura Crespo and Pedro Mira:* “Caregiving to elderly parents and employment status of European mature women”.

1008 *Enrique Moral-Benito*: “Model averaging in economics”.

1009 *Samuel Bentolila, Pierre Cahuc, Juan J. Dolado and Thomas Le Barbanchon:* “Two-tier labor markets in the Great Recession: France vs. Spain”.

1010 *Manuel García-Santana and Josep Pijoan-Mas:* “Small Scale Reservation Laws and the misallocation of talent”.

1101 *Javier Díaz-Giménez and Josep Pijoan-Mas:* “Flat tax reforms: Investment expensing and progressivity”.

1102 *Rafael Repullo and Jesús Saurina:* “The countercyclical capital buffer of Basel III: A critical assessment”.

1103 *Luis García-Álvarez and Richard Luger:* “Dynamic correlations, estimation risk, and portfolio management during the financial crisis”.

1104 *Alicia Barroso and Gerard Llobet:* “Advertising and consumer awareness of new, differentiated products”.

1105 *Anatoli Segura and Javier Suarez:* “Dynamic maturity transformation”.

1106 *Samuel Bentolila, Juan J. Dolado and Juan F. Jimeno:* “Reforming an insider-outsider labor market: The Spanish experience”.

1201 *Dante Amengual, Gabriele Fiorentini and Enrique Sentana:* “Sequential estimation of shape parameters in multivariate dynamic models”.

1202 *Rafael Repullo and Javier Suarez:* “The procyclical effects of bank capital regulation”.

1203 *Anne Layne-Farrar, Gerard Llobet and Jorge Padilla:* “Payments and participation: The incentives to join cooperative standard setting efforts”.

1204 *Manuel Garcia-Santana and Roberto Ramos:* “Dissecting the size distribution of establishments across countries”.