The firm size distribution across countries and skill-biased change in entrepreneurial technology*

Markus Poschke†

McGill University,
CIREQ and IZA

November 2017

Abstract

Development is associated with systematic changes in the firm size distribution. I document that the mean and dispersion of firm size are larger in rich countries, and increased over time for U.S. firms. To analyze the firm size-development link, I construct a frictionless general equilibrium model of occupational choice with skill-biased change in entrepreneurial technology (i.e., technical progress favors better entrepreneurs). The model accounts for key aspects of the U.S. experience with only changes in aggregate technology. It attributes half the variation in mean and dispersion of firm size across countries to technical change. Distortions also affect the size distribution.

JEL codes: E24, J24, L11, L26, O30

Keywords: occupational choice, entrepreneurship, firm size, skill-biased technical change

*I would like to thank Francisco Alvarez-Cuadrado, Alessandra Bonfiglioli, Rui Castro, Russell Cooper, Gino Gancia, George Jia, Chad Jones, Bart Hobijn, Pete Klenow, Fabian Lange, Mariacristina De Nardi, Tapio Palokangas, Nicolas Roys, Ayşegül Şahin, Roberto Samaniego and seminar participants at the Federal Reserve Banks of Chicago, San Francisco, New York and Philadelphia, at the World Bank, Carleton University, McGill University, George Washington University, York University, the University of Barcelona, the Université de Montréal macro brownbag, the XXXIV Simposio de Análisis Económico (Valencia 2009), the Society for Economic Dynamics 2010 Meeting in Montreal, the 7th Meeting of German Economists Abroad (Frankfurt 2010), the Cirpée-Ivey Conference on Macroeconomics and Entrepreneurship (Montréal 2011), the European Economic Review Young Economist Workshop (Bonn 2011), the CESifo Conference on Macroeconomics and Survey Data (Munich 2011), the Canadian Macro Study Group (Vancouver 2011), LACEA (Lima 2012) and the Barcelona GSE Workshop on Firms in the Global Economy (Barcelona 2014) for valuable comments and suggestions, Lori Bowan at the U.S. Census Bureau for providing a detailed tabulation of firm size counts and Steven Hipple at the Bureau of Labor Statistics for providing information on employment by the self-employed. I gratefully acknowledge financial support from the Fonds de la recherche sur la société et la culture (grant 2010-NP-133431) and from the Social Sciences and Humanities Research Council (grant 410-2011-1607).

†Contact: McGill University, Economics Department, 855 Sherbrooke St West, Montreal QC H3A 2T7, Canada. e-mail: markus.poschke@mcgill.ca

1 Introduction

A large literature documents and studies differences in the firm size distribution across countries (see e.g. Tybout (2000), Alfaro, Charlton and Kanczuk (2009) and Bento and Restuccia (2017)). A common view is that the differences observed across rich and poor countries reflect distortions. In this paper, I argue that development is associated with systematic changes in the firm size distribution and that a significant component of observed cross-sectional differences across rich and poor countries is accounted for by differences in the level of development rather than distortions.

To do so, I first systematically document differences in the firm size distribution across rich and poor countries, using data collected in a harmonized way. This analysis yields two new facts: First, the average size of firms is significantly higher in rich countries, with an elasticity of average size with respect to country income per worker in excess of 0.5. Second, firm size is significantly more dispersed in rich countries. These facts hold up in two datasets, covering firms in more than 40 countries and in all sectors except agriculture.[1]

Parallel patterns hold in U.S. history: data from several sources show that the mean and dispersion of firm size there have also increased with development, and that employment has become more concentrated in large firms. Since these changes in the U.S. firm size distribution with development are unlikely to be driven by trends in distortions, they constitute a first indication that at least part of the differences between rich and poor countries may be directly attributable to development.

To pursue this argument further and to be able to analyze it quantitatively, I develop a model that is consistent with these patterns in the data. Given the U.S. experience as a point of reference, this is a frictionless occupational choice model à la Lucas (1978) with two additional features: technological change does not benefit all potential entrepreneurs equally, and an individual's potential payoffs in working and in entrepreneurship are positively related.

I call the first additional feature skill-biased change in entrepreneurial technology (SCET). The idea is that as the menu of available technologies expands, raising aggregate productivity (assuming love of variety, as in Romer 1987), individual firms have to cope with increasing complexity of technology.[2] SCET then means that while advances in the technological

[1] The combination of these facts suggests an outward shift in the right tail of the firm size distribution with development. This is in line with two further patterns from the data: Rich countries have more large firms and a size distribution that is more skewed to the right.

[2] Jovanovic and Rousseau (2008) document that from 1971 to 2006, the average yearly growth rates of the stocks of patents and trademarks in the U.S. were 1.9% and 3.9%, respectively, implying a substantial increase in variety. Michaels (2007) computes an index of complexity based on the variety of occupations employed in an industry. He shows that complexity in U.S. manufacturing has increased substantially over the past century and a half, and that complexity was higher in the U.S. than in Mexico. Similarly, every

frontier give all firms access to a more productive technology, they do not affect all firms equally. Some firms can use a larger fraction of new technologies than others. As a result, some firms remain close to the frontier and use a production process involving many, highly specialized inputs, while others fall behind the frontier, use a simpler production process, and fall behind in terms of relative productivity.[3]

The second crucial assumption is that agents differ in their labor market opportunities, and that more productive workers can also manage more complex technologies if they become entrepreneurs. Occupational choice between employment and entrepreneurship closes the model. Because advances in the technological frontier do not benefit every potential entrepreneur equally, the position of the frontier then governs occupational choice. The more advanced the frontier, the greater the benefit from being able to stay close to it, as other firms fall behind. In equilibrium, advances in the frontier also raise wages, so that entrepreneurs' outside option improves, and marginal entrepreneurs exit. As a consequence, technical change leads high-productivity firms to gradually expand their operations as their productivity improves more than others'. Their entry and growth raise labor demand and the wage, implying that low-productivity entrepreneurs eventually find employment more attractive and exit. This in turn affects the firm size distribution. In particular, average firm size rises with development, and size dispersion increases with development under some conditions.

These results are qualitatively in line with the evolution of the U.S. firm size distribution over the last few decades. To evaluate the quantitative performance of the model, I calibrate it to U.S. data, using both recent cross-sectional data and historical data on average firm size. Parameterized in this way, the model matches the observed increase in average firm size by design, and in addition performs well in terms of the increasing size dispersion and shift towards large firms observed since the 1970s. This shows that a frictionless model with SCET and occupational choice can do a good job at explaining variation in firm size distributions – in this case those observed in the U.S. over the last few decades. The model can thus be taken as a benchmark model for the differences in the firm size distribution that one should expect, in the absence of frictions.

How well can the frictionless model account for differences in the firm size distribution

new classification of occupations in the U.S. from 1970 to 2010 lists more occupations than the preceding one (Scopp 2003).

[3] In line with this, Cummins and Violante (2002) find that the gap between the frontier and average technology in use has been increasing in the U.S. over the entire span of their data (1947-2000), implying that firms have not all benefitted equally from technology improvements. Similarly, Bloom, Sadun and Van Reenen (2012) find that gains from the introduction of information technology differed both across firms and across countries.

across countries? To evaluate this, I compute how the firm size distribution in the model changes with development, using the U.S. calibration and changing only the parameter governing aggregate technology. The model generates elasticities of the mean and dispersion of firm size with respect to output per worker that are almost half of those estimated in the data. Changes in occupational choice in response to SCET are crucial for this result. This suggests that development on its own is responsible for a large fraction of observed differences in the firm size distribution between rich and poor countries.

To complement this analysis, I also explore the potential impact of size- or productivity-dependent distortions à la Restuccia and Rogerson (2008) on the firm size distribution. I model size-dependent distortions (SDDs) in a simple way and assume heavier distortions in poorer countries, as suggested by the literature (see e.g. Hsieh and Klenow 2009). Heavier distortions reduce average firm size and size dispersion. Quantitatively, the model with SCET and SDDs can account for the entire systematic variation in average size with income per worker, but significantly overstates cross-country variation in dispersion. I conclude that both SCET and SDDs are important determinants of differences in firm size distributions, and outline some directions for future work.

Related literature. Several studies have documented aspects of the firm size distribution across countries. An early contribution is Tybout (2000), who surveys the literature and shows that manufacturing employment is more concentrated in large plants in richer countries. Work attempting to expand on this has been hobbled by limited comparability of data across countries. For example, the World Bank Enterprise Surveys used by García-Santana and Ramos (2015) do not cover the informal sector, which is large in poor countries, while the Dun & Bradstreet (D&B) data used by Alfaro et al. (2009) tend to oversample large firms, in particular in poorer countries, where D&B's coverage is thinner. A similar issue affects the United Nations Industrial Development Organization's (UNIDO) Industrial Statistics Database used by Bollard, Klenow and Li (2016). As a result, these sources overstate firm sizes in poor countries.

To overcome this problem, some recent work has used manufacturing censuses that also include small firms. Hsieh and Klenow (2014) do so for three countries, the U.S., Mexico, and India, to compare plant size-age profiles across these countries. In a very recent paper, Bento and Restuccia (2017) also draw on data from manufacturing censuses and similar sources that include small firms to measure mean establishment size in manufacturing for a large number of countries. In line with this paper, they find a strong, positive relationship between development and mean establishment size. Their strategy does not allow examining higher moments of the size distribution, like dispersion. Moreover, it is limited to manufacturing, which accounts for only a fraction of overall employment.

Several papers have advanced theories predicting that average firm size increases with development (see e.g. Lucas (1978), Gollin (2007), Akyol and Athreya (2009) and Roys and Seshadri (2013)). However, none of them predicts the observed increase in size dispersion with development. Bento and Restuccia (2017) and Hsieh and Klenow (2014) explore the effect of distortions on mean plant size and life cycle plant growth, respectively, but do not consider their effect on size dispersion. To my knowledge, no paper has analyzed the quantitative effect of distortions on the firm size distribution for a broad cross-section of countries.[4]

The paper is organized as follows. Section 2 describes the data and documents relevant facts about entrepreneurship and the firm size distribution. Section 3 presents the model, and Section 4 shows how entrepreneurship and characteristics of the firm size distribution change with development. Section 5 presents quantitative results for the benchmark model, both for U.S. history and for the cross-section of countries. Section 6 explores the effect of size-dependent distortions, and Section 7 concludes.

2 Entrepreneurship, the firm size distribution and development

In this section, I show facts on the firm size distribution across countries using two complementary data sets. Obtaining data on the firm size distribution across countries is notoriously hard because measurement in national surveys or administrative data is not harmonized across countries. The Global Entrepreneurship Monitor (GEM) and the Amadeus data base collected by Bureau Van Dijk constitute two exceptions.[5] To the best of my knowledge, this is the first paper using GEM data across countries for general equilibrium analysis, and one of the first to use Amadeus for this purpose. I next describe these sources, and show that when compared to available administrative data, both give an accurate representation of the bulk of the firm size distribution, with the exception of only its right tail in the GEM and

[4] Other work on entrepreneurial choice and development has typically focussed on the role of credit constraints, see e.g. Banerjee and Newman (1993), Lloyd-Ellis and Bernhardt (2000) and Akyol and Athreya (2009). The first two of these papers focus on the role of the wealth distribution when there are credit constraints. The latter in addition takes into account how the outside option of employment varies with income per worker, and how this affects entrepreneurial choice.

[5] Another exception are some OECD publications such as Bartelsman, Haltiwanger and Scarpetta (2004) or Berlingieri, Blanchenay and Criscuolo (2017) that provide information on some OECD countries and a limited number of other countries. Their numbers arise from an effort to process national official data to make it comparable, while in the case of the GEM and Amadeus, data collection is already harmonized.

its left tail in Amadeus. In the core of this section, I then use the two data sets to show that the mean and dispersion of firm size are significantly larger in richer countries, and that they increased over time in U.S. history.

2.1 Data sources

2.1.1 The Global Entrepreneurship Monitor (GEM) survey

The GEM is an individual-level survey run by London Business School and Babson College now conducted in more than 50 countries. Country coverage has been expanding since its inception in 1999, with data for several years available for most countries. The micro data is in the public domain, downloadable at `http://www.gemconsortium.org/`. Most developed economies are represented, plus a substantial number of transition and developing economies, ensuring that the data covers a wide variety of income levels.[6]

The survey focusses on entrepreneurship. That is, while the survey overall is conducted by local research organizations or market research firms to be representative of a country's population, it contains only limited demographic information (e.g. education) on non-entrepreneurs. It contains much richer information on entrepreneurs, including their firm's employment.

Importantly, the survey is designed to obtain harmonized data across countries. It is thus built to allow cross-country comparisons, the purpose for which it is used here. In addition, because it is an individual-level survey, it captures all types of firms and not just firms in the formal sector or above some size threshold. For studying occupational choice, this is evidently important. This feature makes the GEM data a valuable source of information for the purposes of the analysis in this paper, and differentiates it from firm- or establishment-level surveys such as the World Bank's Entrepreneurship Survey, which covers only registered firms. Moreover, Reynolds et al. (2005), Acs, Desai and Klapper (2008) and Ardagna and Lusardi (2009) have shown that patterns found in GEM data align well with those based on other sources. The main weakness of the GEM as a source of information on firms is that, because it is a household survey, publicly listed firms with dispersed ownership are not included. The use of data from Amadeus addresses this issue.

Country averages for some measures are easily available on the GEM website. In the following, I use the underlying micro data for the years 2001 to 2005 to obtain statistics on the firm size distribution, for which no country-level numbers are reported. For this period, data is available for 44 countries, though not for all years for all countries. Pooling

[6] Inclusion in the survey depends on an organization within a country expressing interesting and financing data collection. For a list of countries in the sample, see Table 9 in the Appendix.

the available years for each country, the number of observations per country is between 2,000 in some developing economies and almost 80,000 in the UK, with a cross-country average of 11,700. This is sufficient for computing the summary statistics of the firm size distribution that I use in the following. Unfortunately, in many countries, there are not enough observations for obtaining reliable estimates for more detailed size classes, so I rely on summary statistics for the entire distribution.[7]

The GEM captures different stages of entrepreneurial activity. I consider someone an entrepreneur, and include the firm in the analysis, if they declare running a firm that they own and they have already paid wages (possibly to themselves, for the self-employed). I then obtain firm size data for these firms.

2.1.2 Amadeus

This database contains financial and employment information on more than five million companies from 34 European countries, including all of the European Union. The data is collected by the company Bureau Van Dijk (BvD). BvD and its local subsidiaries collect data on public and private companies, which under European regulations typically are required to file some financial information in publicly accessible local registers. The information in Amadeus thus stems from companies' official filed and audited accounts, with the exception of data for some Eastern European countries, which is collected from the companies themselves. The Amadeus database has been used for firm level analysis by Bloom et al. (2012), among others.

The Amadeus data covers all sectors, except for banks and insurance companies. In Western Europe, according to information provided by BvD, firms with at least 300 employees are typically required to publicly file their accounts. The threshold can be lower in some countries, and can also depend on a firm's legal status. Companies with a legal status conferring limited liability typically are required to file. As shown next, coverage is excellent for firms with more than 250 employees, and very good even for medium-sized firms with 50 to 249 employees.

In the following, I use Amadeus data from the 33 countries where total employment in the database corresponds to at least 10% of private sector employment in the country. I use data for 2007, to exclude the effect of the deep global recession in the following years.[8]

[7] I use data for all countries except for Latvia, for which mean employment is 60% above the next-highest country value.

[8] Private sector employment is computed as total employment from the World Development Indicators minus general government employment, from the same source. Qualitative results are not sensitive to using different coverage cutoffs, like 0.33 or 0.5, though of course significance of results suffers from dropping observations. Results are also similar when all years are used. See Table 9 for a list of countries.

2.1.3 How well do data from the Global Entrepreneurship Monitor and Amadeus reflect the global firm size distribution?

The differential focus of the GEM and Amadeus may raise concerns of representativeness. Since the GEM focusses on small firms and Amadeus on larger ones, it is possible that each provides a good picture of part of the firm size distribution, while missing the middle of the distribution.

By comparing GEM and Amadeus data to the comparable administrative data that is available, this section reveals that this concern is not warranted. Comparison data come from Eurostat, the Statistics of U.S. Businesses (SUSB), and INSEE, the French national statistical institute.

Eurostat provides data on the firm size distribution for a very limited number of size classes in its Structural Business Statistics (SBS), drawing on both surveys and administrative sources.[9] Table 1 shows a comparison of the firm size distribution in Eurostat to that computed from the GEM, averaging across the twelve countries with comparable data.

The table shows that the GEM information on the distribution of medium-sized firms (10-249 employees) is excellent. The GEM is weaker in its coverage of the extremes of the size distribution. Its coverage of large firms (more than 250 employees), for which it was not designed, does not appear all too reliable. On average, it also understates the share of small firms (less than 10 employees). Importantly for the purposes of this paper, the GEM/Eurostat discrepancy at the country level is unrelated to GDP per worker. (Regressing the ratio of small-firm shares from the GEM and Eurostat on log GDP per worker outside agriculture results in a coefficient of 0.007, with a standard error of 0.04.) It should thus not affect the conclusions drawn in the remainder of the section. In Section 2.2, I also show more formally that the facts established there are robust to the influence of the 0-9 size category.[10,11]

Turning to large firms, Figure 1(a) shows the number of firms with 250 or more employees in Amadeus relative to that in Eurostat for the 19 countries represented in both data sets.

[9] See `http://ec.europa.eu/eurostat/cache/metadata/en/sbs_esms.htm` for a detailed description. I use 2010 data to maximize country coverage. Size categories are 0-9, 10-19, 20-49, 50-249 and 250+ employees. This rough classification makes it hard to gain insights on the firm size distribution across countries from the Eurostat data on its own.

[10] For the U.S., the GEM slightly overstates the share of small firms compared to SUSB data: by two percentage points for firms with 0-19 employees, and 6 percentage points for 0-9 employees. For larger size groups, the two sources accord very closely.

[11] Where does the GEM/Eurostat discrepancy come from? A first potential explanation is a stricter sample inclusion criterion for my GEM sample compared to Eurostat. This is possible, since SBS data may also include some inactive companies. A second potential explanation is that, unlike Eurostat, the GEM data do not reveal whether individuals own multiple businesses.

Table 1: Firm size distribution, GEM versus Eurostat, average across 12 countries

share (%) of firms with...	Eurostat	GEM
< 10 employees	93.3	86.2
10-249 employees	6.5	12.8
among these:		
10-19 employees	54.2	50.0
20-49 employees	30.9	34.0
50-249 employees	14.9	15.9
250 and more employees	0.2	1.0

Note: Figures are arithmetic averages of the data for Austria, Belgium, Spain, France, Croatia, Hungary, the Netherlands, Norway, Poland, Portugal, Sweden and Slovenia. Eurostat data is for 2010.

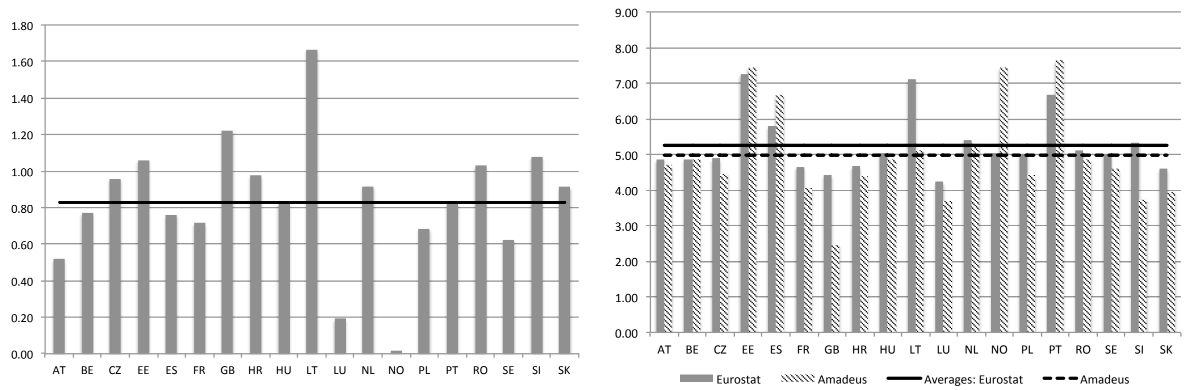
A value of one indicates that the number of large firms is identical in the two datasets. The horizontal line shows the average of the ratio across countries. The figure shows that on average, the ratio is close to one, implying that Amadeus does an excellent job at capturing the universe of large firms.

Figure 1(b) shows the ratio of the number of firms with 50 to 249 employees to that with 250 or more employees, for both Amadeus and Eurostat. Again, the horizontal lines show cross-country averages. It is clear that on average, the ratios are very close, with about five medium-sized firms for each large firm in the 19 countries under consideration in both sources. More than this, the ratios computed from the two data sets are extremely close even for many individual countries. This implies that the shares of large and medium sized firms among firms with 50 or more employees are very close in Amadeus and Eurostat data. Amadeus thus provides an excellent picture of the size distribution of firms with 50 or more employees across the 19 countries under consideration.[12]

Finally, I compare the firm size distribution in Amadeus to that in French official data using the extremely detailed data on firm sizes for France reported in Gourio and Roys (2014).[13] Figure 2 compares the size distribution for firms with twenty to one hundred

[12]Just as for the GEM data, differences between the Eurostat and Amadeus data most likely are due to differences in the definition of a firm. Amadeus aims to attribute firm information to the ultimate owner. Eurostat in contrast defines a firm as "the smallest combination of legal units that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources." As a result, Amadeus may overall feature a smaller number of larger, more consolidated firms. This also implies that for some large size categories, the number of firms in Amadeus could exceed that in Eurostat.

[13]Thanks to Nicolas Roys for providing the detailed firm size counts underlying Figure 1 in Gourio and Roys (2014). The data is complied by INSEE, the French statistical institute, based on a combination of administrative data and surveys. The data Gourio and Roys use is for the years 1994-2000.



(a) Number of firms with more than 250 employees, Amadeus/Eurostat

(b) Number of medium (50-249 employees) relative to large (>250) firms, Amadeus and Eurostat

Figure 1: Comparison of Amadeus and Eurostat data

Notes: Horizontal lines indicate averages. Amadeus data is for 2007. Eurostat data is series sbs_sc_sca_r2 for the year 2010. See `http://ec.europa.eu/eurostat/cache/metadata/en/sbs_esms.htm` for a detailed description, and Table 9 for country codes.

employees. Because Gourio and Roys aggregate data across years, the figure plots the ratio of the number of firms at each integer level of employment from 20 to 100 to the number of firms with employment of exactly 100 for both sources. Close agreement of the two sources is obvious, including even the kink in the distribution at 50 employees that is the focus of Gourio and Roys (2014) and Garicano, Lelarge and Van Reenen (2017). Below 50 employees, the accuracy of the size distribution obtained from Amadeus gradually deteriorates. The fit is similarly good for larger firms. In the INSEE data, the ratio of the number of firms with more than 200 (1000) employees to that with between 50 and 100 employees is 0.57 (0.074). In Amadeus, this figure is 0.52 (0.081).

To summarize, the GEM provides a reliable picture of the firm size distribution for firms with less than 250 employees, while slightly understating the share of very small firms (0 to 9 employees). Amadeus data capture a very large fraction of firms with 250 or more employees, and provide an excellent picture of the size distribution of firms with 50 or more employees. These statements are based on cross-country averages, and on comparisons of broad size classes. Sampling error may lead to differences for individual countries, or for narrower size classes. Each of the two data sets thus provides a reliable image of part of the size distribution, with a substantial area of overlap in which both do a good job.

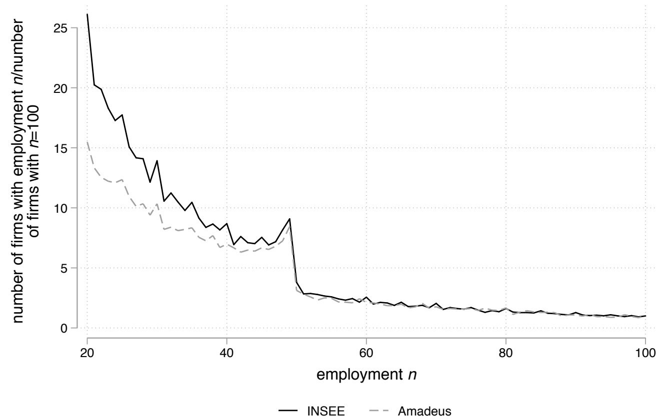


Figure 2: Number of firms of size n relative to number of firms of size 100, data for France from INSEE and from Amadeus.

Source for the INSEE data: Gourio and Roys (2014).

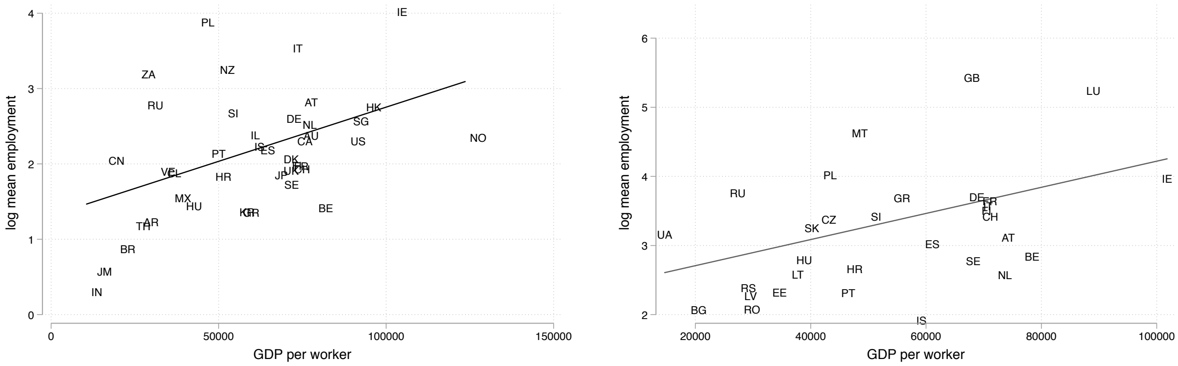
2.2 Cross-country evidence

Next, I use the GEM and Amadeus data to establish two new facts on the firm size distribution and income per worker. To do so, I show plots of moments of the firm size distribution against 2005 PPP GDP per worker outside agriculture in Figures 3 to 5.[14] Results are generally similar whether using the level or log of GDP per worker. Each figure also contains an OLS line of best fit. The regression lines drawn in the figures are all significant at least at the 5% level. Tables 2 and 3 report bivariate regression results using the log of GDP per worker outside agriculture. They also contain measures of fit, which are high for a bivariate relationship in cross-sectional data.

Fact 1. *Average firm employment increases with income per worker (see Figure 3).*

It is clear from Figure 3 that average firm employment is larger in richer countries in both the GEM and the Amadeus data. Regression results shown in Table 2 show that the elasticity of average employment with respect to income per worker is around 0.75-0.8 in both data sources. Regression results in Table 3 show that the positive relationship persists

[14]By its sampling procedure, the GEM captures few agricultural businesses (only 4% on average). Accordingly, the model described below should be interpreted as referring to the non-agricultural parts of the economies studied. In line with this, I combine data on real GDP and persons engaged from the Penn World Tables 8 with information on value added and persons engaged in agriculture from the FAO to compute output per worker outside agriculture. (See Heston, Summers and Aten (2009) and Feenstra, Inklaar and Timmer (2015) for background, and `http://www.rug.nl/ggdc/productivity/pwt/pwt-releases/pwt8.0` and `http://faostat.fao.org` for the data.)



(a) GEM data

(b) Amadeus data

Figure 3: Average employment and income per worker.

Notes: GDP per worker outside agriculture is computed as real GPD for 2005 at purchasing power parity from the Penn World Tables 8 (Summers and Heston 1991, Heston et al. 2009) minus value added in agriculture, forestry and fishing (from FAO macro indicators), divided by total persons engaged minus persons engaged in agriculture, also from the FAO. Firm employment data from the GEM for the left panel and from Amadeus for the right panel. The vertical axis shows log average employment. The lines represent the best linear fits. Regression results are reported in Table 2.

when using only data for the part of the firm size distribution for which each dataset is most reliable. Specifically, the relationship between average size and income per worker is significantly positive also when excluding the self-employed, firms with fewer than ten employees, or firms with more than 250 employees in the GEM, and when excluding firms with fewer than 250 employees in Amadeus.

The differences in coefficients between Tables 2 and 3 are due to the fact that not only average size, but also the importance of large firms is greater in richer countries. (See Figure 11 and Table 10 in the Appendix, which show that the fraction of firms with more than 10 employees and the fraction of employment in firms with more than 250 employees is greater in richer countries.) This implies that excluding small or large firms from the analysis, as done in the robustness checks, makes firm sizes more similar across countries with different income levels. As a consequence, regression coefficients are lower for the truncated samples.[15]

To summarize, both data sources show a clear, strong positive relationship between average firm size and income per worker, no matter whether all data is used or whether the samples are truncated.

[15]In addition, Table 11 in the Appendix shows that results are similar when U.S. sector weights for manufacturing and services are used in the regressions, ruling out the potential influence of structural change.

Table 2: The firm size distribution and income per worker.

	GEM data			Amadeus data		
Moment	Coeff.	SE	R^2	Coeff.	SE	R^2
Log average employment	0.718***	(0.191)	0.266	0.824**	(0.325)	0.182
Entrepreneurship rate	-0.040***	(0.014)	0.174			
Standard deviation of log employment	0.228***	(0.055)	0.308	0.183**	(0.089)	0.131

Notes: Data sources as in Figure 3. The table shows coefficients from bivariate regressions of each moment on log GDP per worker outside agriculture, the standard errors on those coefficients, and the R^2 for each regression. A constant is also included in each regression (coefficient not reported). The regression for the log standard deviation using Amadeus data excludes Ukraine, which is an outlier here. (This is visible in Figure 5; results are qualitatively similar when including it.) The preceding figures show these relationships for the level instead of the log of GDP. *** (**) [*] denotes statistical significance at the 1% (5%) [10%] level.

Table 3: The firm size distribution and income per worker – robustness checks.

Moment	Coeff.	SE	R^2	Coeff.	SE	R^2
	GEM data (excl. self-employed)			*GEM data ($n \geq 10$)*		
Log average employment	0.653***	(0.179)	0.255	0.417**	(0.183)	0.118
Standard deviation of log employment	0.228***	(0.055)	0.308	0.190**	(0.087)	0.109
	GEM data ($n < 250$)			*Amadeus data ($n \geq 250$)*		
Log average employment	0.530***	(0.097)	0.435	0.537***	(0.144)	0.323
Standard deviation of log employment	0.172***	(0.029)	0.480	0.196***	(0.054)	0.308

Notes: Data sources as in Figure 3. Other remarks as in Table 2.

Similar patterns had previously been documented for a more limited number of countries. For instance, Hsieh and Klenow (2014) show that U.S. firms are larger than Mexican and Indian ones. Tybout (2000) and references therein show that small firms account for a much larger share of employment in poorer countries. Finally, more recently, Bento and Restuccia (2017) have shown that establishments in the manufacturing sector are systematically larger in richer countries. They find an elasticity of size with respect to GDP per capita of 0.35, close to that shown in Table 3. The evidence presented here allows extending results from these papers to a much larger number of countries and beyond the manufacturing sector. The latter is important, given the limited importance of manufacturing in rich economies (for example, manufacturing value added has accounted for less than 20% of U.S. GDP since the 1970s) and the well-known differences in scale across sectors (see e.g. Buera and

Kaboski 2012).

The larger size of firms in richer countries relates to another systematic pattern in the data, namely the finding by Gollin (2007) that the self-employment rate falls with income per capita in ILO data. Figure 4 and Table 2 show that this relationship is strongly reproduced in GEM data. The pattern may appear to contrast with some publications that report a larger number of firms or establishments per capita in richer countries (see e.g. Alfaro et al. 2009, Klapper, Amit and Guillén 2010). The reason for that is that the population of entrepreneurs under consideration matters. The pattern shown in Figure 4 holds for broad measures of entrepreneurship that include small firms, and not just large or registered ones. Such a broad measure is the appropriate one for studying occupational choice, which requires considering all types and sizes of firms. Contrasting results in other sources can be attributed to the use of sources that miss small firms, and are more likely to do so in poorer countries. Because of high rates of informality in poor countries, this is the case with data based on business registries.[16]

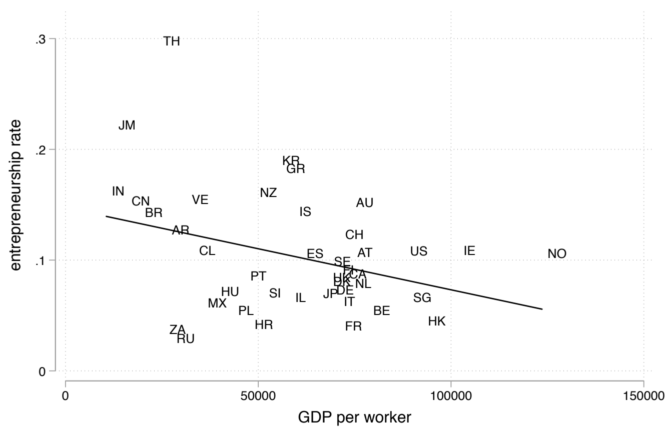


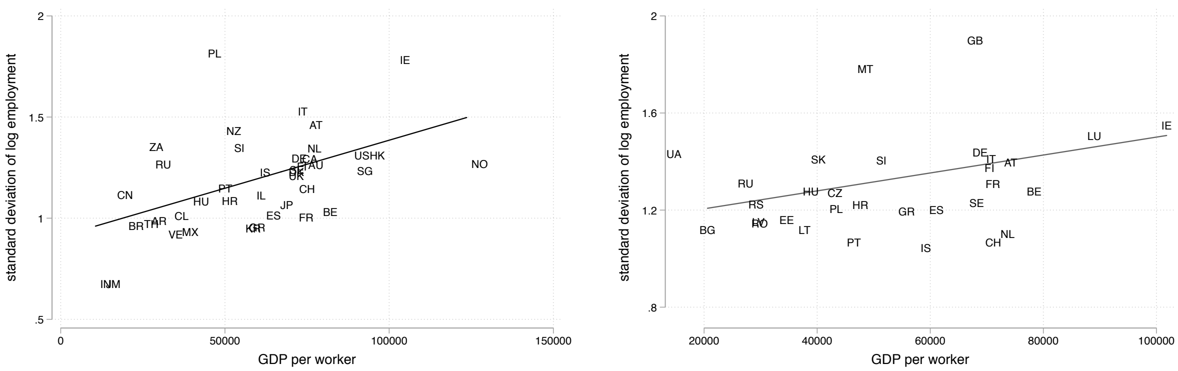
Figure 4: The entrepreneurship rate and income per worker.

Notes: The entrepreneurship rate is computed from GEM data. Entrepreneurs are defined as survey respondents who declare running a firm that they own and who have already paid wages, possibly to themselves. Other sources and further remarks as in Figure 3.

Fact 2. *The dispersion of firm size in terms of employment increases with income per worker (see Figure 5).*

[16]As discussed above, the GEM data suffer from the opposite problem, since they miss firms that are not privately owned, and undersample large firms ($n > 250$). Given the small number of large firms, this is unlikely to affect the overall pattern. For example, firms with more than 100 employees account for only about 0.4% of all firms in U.S. Census SUSB data.

Figure 5 shows a clear positive relationship between the standard deviation of log firm size and income per worker. The relationship is very similar for other measures of dispersion, like the interquartile ratio. Regression results in Tables 2 and 3 show that the pattern is statistically significant and robust. The only previous mention of such a relationship I found is Bartelsman et al. (2004), who show that firm size dispersion is substantially higher in industrialized countries compared to emerging markets, using OECD and World Bank data for a much smaller set of countries.[17]



(a) Small and medium sized firms (GEM, $n < 250$)

(b) Large firms (Amadeus, $n \geq 250$)

Figure 5: Standard deviation of log employment and income per worker.

Notes: Data sources and further remarks as in Figure 3.

The finding of varying dispersion is important, because it indicates that larger average size in richer countries is not simply due to a shift to the right of the firm size distribution. Instead, the combination of higher mean size and higher dispersion in richer countries indicates an outward shift in the right part of the distribution.

For a single-peaked, right-skewed distribution like the firm size distribution, higher dispersion will typically go along with higher skewness. Figure 12 and Table 10 in the Appendix show that skewness, measured in a way that is robust to outliers, is indeed higher in richer countries.

[17]The well-known paper Hsieh and Klenow (2009) shows larger TFP dispersion across manufacturing establishments in China and India compared to the U.S.. Variation across these countries in the lower size threshold for sample inclusion implies that these results are not comparable to the ones here.

2.3 The firm size distribution in U.S. history

Differences in the firm size distribution with development can be studied across countries, or within a country over time. This Section documents historical trends in the U.S. size distribution, using a variety of sources.

One of the few references on trends in the firm size distribution in the U.S. is the seminal paper by Lucas (1978), who reported that average firm size increased with per capita income over U.S. history (1900-70). Figure 6 shows that this time-series relationship persists. It reports measures of average firm size close to those used by Lucas (the two series labelled "BEA Survey of Current Business" and "Dun & Bradstreet", both from Carter et al. 2006) and more recent data. The most recent available series is from U.S. Census Business Dynamics Statistics (BDS). It covers employer firms accounting for 98% of U.S. private employment. In order to obtain average firm size for a broader measure of firms I also report average firm size when taking into account non-employer firms, or self-employed without employees. This measure is obtained by combining BDS data with data on unincorporated self-employed businesses reported in Hipple (2010).[18] While the five series shown in Figure 6 cover slightly different populations of firms, they all show an increasing trend, except for the interwar period. This upward trend of course occurs simultaneously with increasing per capita income.[19] Average firm size thus increases with per capita income both in U.S. history and across countries.

BDS data can also be used to assess the evolution of other moments of the size distribution since the late 1970s. Table 4 shows that the fraction of small firms and their share of overall employment declined over this time period. Since the size bins used for organizing BDS data are consistent over time, the data can also be used to assess trends in dispersion. Using size bin means, the standard deviation of log employment increased by 0.076 from 1977 to 2008. It increased by 0.067 when size bin midpoints are used.

Other sources suggest a similar pattern. Using the U.S. Census of Manufactures, Bonfiglioli, Crinó and Gancia (2015) find an increase in the standard deviation of log sales among U.S. manufacturing plants in the period 1997 to 2007. Autor, Dorn, Katz, Patterson and

[18]Unfortunately, this series is rather short. This is because information on employment by the unincorporated self-employed is only available starting in 1995. Many thanks to Steven Hipple for providing some additional information.

[19]According to historical U.S. manufacturing census data, this process started even earlier. Using data from the Atack and Bateman (1999) national samples of manufacturing establishments, Margo (2013, Table 1) reports that average firm size in U.S. manufacturing increased by 46% over the period 1850 to 1880. In fact, already Atack (1986) drew attention to the fact that over the course of the 19th century, large firms expanded in U.S. manufacturing, but small firms persisted, while losing market share.

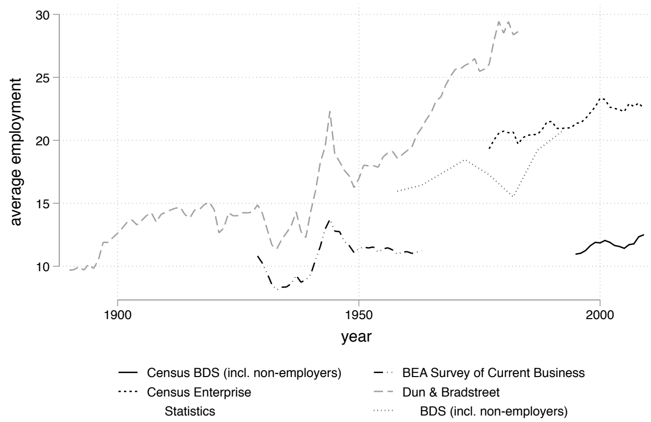


Figure 6: Average firm size (employment) over U.S. history, 1890-2009

Sources: Census Bureau Business Dynamics Statistics (BDS): data available at `http://www.ces.census.gov/index.php/bds`; when including non-employers, combined with Current Population Survey (CPS) data reported in Hipple (2010); Census Enterprise Statistics series: from various Census reports; BEA Survey of Current Business series: from Carter et al. (2006, Series Ch265); Dun & Bradstreet series: from Carter et al. (2006, Series Ch408). The first three sources also report total employment. For the last two series, employment is from Carter et al. (2006, Series Ba471-473 and Ba477). The Dun & Bradstreet firm counts exclude finance, railroads and amusements. Adjusting employment for this using Series Ba662, Dh31, Dh35, Dh53 and Df 1002 shortens the series without affecting the trend. Starting 1984, Dun & Bradstreet gradually cover additional sectors, at the cost of comparability over time, so I only plot data up to 1983. Series Ch1 in Carter et al. (2006), which draws on Internal Revenue Service data, also contains historical firm counts but is less useful because of frequent changes of definition, in particular for proprietorships. BDS data is aggregated annual data based on the Longitudinal Business Database (LBD) maintained by the Census Bureau's Center for Economic Studies which draws on, among other sources, the Business Register, Economic Censuses and IRS payroll tax records.

Van Reenen (2017) also use Census data and show increasing concentration of employment and sales at the top within four digit industries over the last few decades. Kehrig (2012) shows using data from the Annual Survey of Manufactures that the standard deviation of plant-level total factor productivity (TFP) in the U.S. has increased by about 52% between 1977 and 2006. (This excludes a further upward jump in the deep recession in the following years.)[20] Finally, Elsby, Hobijn and Şahin (2013) show that there have been large increases in income inequality among proprietors, driven mainly by increases at the top.

2.4 Time-series evidence from other countries

Limited evidence on trends in average firm size and size dispersion is available from other countries. Tomlin and Fung (2012) report that average firm size in Canada increased between 1988 and 1997. Felbermayr, Impullitti and Prat (2013, Table 4) show the same for Germany

[20]Thanks to Matthias Kehrig, Gino Gancia and Alessandra Bonfiglioli for providing this information.

Table 4: Changes in the U.S. firm size distribution, 1979-2007 (5-year averages)

	firms with employment above				
Change in group ... (in percentage points)	5	50	100	250	500
employment share	1.15	3.56	3.63	3.03	2.43
share of firms	1.63	0.67	0.39	0.15	0.07

Notes: Data from the Census BDS. The data contain counts of firms and employment in 12 employment size classes, for employer firms only. Numbers shown are the differences between average shares for the years 2005-2009 and the years 1977-1981. Results are very similar when including more recent data from the recovery from the latest recession.

between 1996 and 2007. A special issue of *Small Business Economics* reveals that average firm size also increased with development over time in several East Asian economies. This is the case in Indonesia (Berry, Rodriguez and Sandee 2002), Japan (Urata and Kawai 2002), South Korea (Nugent and Yhee 2002) and Thailand (Wiboonchutikula 2002). Only in Taiwan, the smallest of these countries, did it fall (Aw 2002). The upward trend in average firm size thus has occurred in a substantial number of countries.

Evidence for other countries also shows increasing dispersion over time, paralleling the evolution in the U.S.: Faggio, Salvanes and Van Reenen (2010) show increasing TFP dispersion for the United Kingdom and Felbermayr et al. (2013) for Germany, while Berlingieri et al. (2017) find it within sectors for a broad set of OECD countries.

This section has shown that the mean and dispersion of firm size are higher in richer countries, and have increased over time in the U.S. and several other countries. The long-running trend in moments of the firm size distribution in the U.S. suggests technological factors as the driving force of changes in the distribution with development. In particular, it seems implausible that time-series differences in the U.S. are driven by a trend in distortions. Accordingly, it seems likely that at least part of the cross-country differences in the firm size distribution should not be attributed to variation in distortions, but to variation in technology across countries that mirrors the development of technology within a country over time. In the following sections, I build and quantitatively evaluate a model to study this possibility, and return to the possible role of distortions at the end of the paper.

3 A simple model

In this section, I present a simple general equilibrium model of occupational choice with skill-biased change in entrepreneurial technology that allows for a transparent analysis of the key economic forces that can generate the facts presented in the previous section. For the quantitative analysis in Section 5, the model will be generalized in a few dimensions.

The economy consists of a unit continuum of agents and an endogenous measure of firms. Agents differ in their endowment of effective units of labor $a \in [0, \bar{a}]$ that they can rent to firms in a competitive labor market. Refer to this endowment as "ability". Differences in ability can be thought of as skill differences. They are observable, and the distribution of ability in the population can be described by a *pdf* $\phi(a)$.

Agents value consumption c of a homogeneous good, which is also used as the numéraire. They choose between work and entrepreneurship to maximize consumption. The outcome of this choice endogenously determines the measures of workers and of firms in the economy.

Labor supply and wage income. Consumption maximization implies that individuals who choose to be workers supply their entire labor endowment. Denoting the wage rate per effective unit of labor by w, a worker's labor income then is wa.

Labor demand and firm profits. Firms use labor in differentiated activities to produce the homogeneous consumption good. They differ in their level of technology M_i, which indicates the number of differentiated activities undertaken in firm i. It thus corresponds to the complexity of a firm's production process, or the extent of division of labor in the firm. A firm's level of technology depends on the entrepreneur's skill in a way detailed below.

A firm's production technology is summarized by the production function

$$y_i = X_i^{\gamma}, \quad X_i = \left(\int_0^{M_i} n_{ij}^{\frac{\sigma-1}{\sigma}} \mathrm{d}j \right)^{\frac{\sigma}{\sigma-1}}, \quad \gamma \in (0,1), \sigma > 1, \tag{1}$$

where y_i is output of firm i, and X_i is an aggregate of the differentiated labor inputs n_{ij} it uses. The production function exhibits decreasing returns to scale. This can be interpreted to reflect any entrepreneur's limited span of control, as in Lucas (1978). It also ensures that firm size is determinate, implying a firm size distribution given any distribution of M_i over firms. The elasticity of substitution among inputs is given by σ. Given that M_i differs across firms and that thus not all firms use all types of differentiated inputs, it is natural to assume that different inputs are gross substitutes ($\sigma > 1$). Heterogeneity in M_i plays a role as long as they are imperfect substitutes. Importantly, the production function exhibits love of variety, and firms with larger M_i are more productive.

The firm's profit maximization problem can be solved using a typical two-stage approach: choose inputs n_{ij} to minimize the cost of attaining a given level of the input aggregate X_i, and then choose X_i to maximize profit. The solution to the latter will depend on a firm's productivity M_i. Dropping firm subscripts, denoting desired output by $\bar{y}$, and defining $\bar{X} = \bar{y}^{1/\gamma}$, the solution to the cost minimization problem yields the firm's labor demand function for each activity j as $n_j(M) = (w/\lambda(M))^{-\sigma} \bar{X}$ for all j, where λ is the marginal cost of another unit of X. With constant returns to scale for transforming the differentiated labor inputs into X, λ is independent of X and equals $M^{\frac{1}{1-\sigma}} w$. Then the demand for each n_j is $n_j(M) = M^{\frac{-\sigma}{\sigma-1}} \bar{X}$ for all j. Because of greater specialization in firms using more complex technologies, their marginal cost of X is lower. As a consequence, they require less of each input to produce $\bar{y}$. Since this implies that M maps one-to-one with TFP, I will refer to it as the firm's productivity.

Choice of $\bar{X}$ to maximize profits yields optimal output and profits as

$$y(M) = \left(\frac{w}{\gamma}\right)^{\frac{-\gamma}{1-\gamma}} M^{\frac{1}{\sigma-1}\frac{\gamma}{1-\gamma}}, \quad \pi(M) = (1-\gamma)y(M). \tag{2}$$

Both output and profits increase in M. They are convex in M if $\gamma > \frac{\sigma-1}{\sigma}$. As this inequality holds for reasonable sets of parameter values (e.g. $\gamma = 0.9$ and $\sigma < 10$), assume that it holds.

Skills and technology. Entrepreneurs run firms and collect their firm's profits. The crucial activity involved in running a firm is setting up and overseeing a technology involving M differentiated activities. Agents differ in their skill in doing this.

To capture this, suppose that an entrepreneur's time endowment is fixed at 1, and that overseeing an activity takes $c(a, \bar{M})$ units of time, where $\bar{M} \geq 1$ is a measure of aggregate technology. Since profits increase in M, each entrepreneur chooses to oversee as many activities as possible given limited time. This implies that $M(a, \bar{M}) = 1/c(a, \bar{M})$. Suppose that the function $M(\cdot)$ satisfies the following five assumptions:

Assumption.

i) $\partial M(a, \bar{M})/\partial a > 0$.

ii) $\partial M(a, \bar{M})/\partial \bar{M} > 0$.

iii) The elasticity of $M(a, \bar{M})$ *with respect to* $\bar{M}$ *is independent of the level of* $\bar{M}$.

iv) The elasticity of $M(a, \bar{M})$ *with respect to* $\bar{M}$ *increases in* a.

v) The elasticity of $M(a, \bar{M})$ with respect to $\bar{M}$ is weakly convex in a.

The first assumption implies that more able individuals can manage more complex production processes and thus run more productive firms. The second assumption implies that, conditional on an entrepreneur's skill, any firm is more productive when situated in a technologically more advanced economy. This allows $\bar{M}$ to drive aggregate output growth. The third assumption helps tractability and is in line with how the effect of aggregate technology on individual firm productivity is typically modelled (see also below). The fourth assumption introduces "skill-biased change in entrepreneurial technology" (SCET): It captures that, while all entrepreneurs benefit from improvements in aggregate technology $\bar{M}$, more skilled entrepreneurs benefit more. Finally, the fifth assumption ensures that profits are convex not only in M, but also in a, and is crucial for the occupational choice patterns discussed below.[21]

Functions fulfilling these assumptions are of the form $\kappa \bar{M}^{\mu(a)}$, where κ is an arbitrary constant, and $\mu(a)$ is positive, increasing and weakly convex in a, and independent of $\bar{M}$. $M(a, \bar{M})$ then is increasing in a and in $\bar{M}$, and its elasticity with respect to $\bar{M}$ is $\mu(a)$. Note also that the assumptions imply that even the least able entrepreneurs can operate at a strictly positive scale ($M(0, \bar{M}) > 0$).

A useful analogy to the existing literature can be drawn for the simplest such function, $\bar{M}^a$. This function is similar to the one often chosen for the marginal cost of innovation in the literature on endogenous growth with R&D. The presence of a in the exponent is akin to introducing heterogeneity in the parameter that controls how existing knowledge affects the productivity of R&D in e.g. Jones (1995).[22] More skilled entrepreneurs are better at drawing on existing knowledge. They are better at exploiting similarities and synergies between different activities, therefore can oversee more of them, and are more productive. As technology advances, the potential for exploiting synergies grows, and more skilled entrepreneurs benefit more from the new technologies.

Another way of interpreting SCET is in relation to the work summarized in Garicano and Rossi-Hansberg (2015). These authors show how declining coordination and communication costs affect the optimal organization of the firms, and allow better managers to run larger

[21] On a technical level, assumption *iv)* is satisfied if M is log supermodular. Chen (2014) makes a similar assumption in a different context. Assumption *v)* is stronger than necessary; what is key for the analysis of occupational choice below is that $M(a)^{\frac{1}{\sigma-1}\frac{\gamma}{1-\gamma}}$ is strictly convex in a.

[22] In that paper, the marginal cost of a unit of knowledge is proportional to $A^{-\phi}$, where A is existing knowledge and ϕ governs the contribution of A to new knowledge creation. The profit function resulting if $M(a, \bar{M}) = \bar{M}^a$ is also closely related to that in the multi-sector model in Murphy, Shleifer and Vishny (1991). There, more able entrepreneurs select into a sector where profits are more elastic with respect to their talent. Differently from here, however, Murphy et al. (1991) assume that aggregate productivity affects all firms' profits equally.

firms. More broadly, one can think that any improvement in technology, be it in management or production technology, allows for new coordination opportunities within the firm. SCET then amounts to assuming that better entrepreneurs gain more from these new opportunities.

Under these assumptions on M, the most able entrepreneurs ($a = \bar{a}$) operate at the technological frontier, the least able ones ($a = 0$) at the lowest level, and intermediate ones at some distance to the frontier. Crucially, for low levels of the frontier, all firms are close to it. The higher the frontier, the more dispersed the levels of technology of potential firms. The actual distribution of technology among active firms depends on occupational choice.

Occupational choice. Occupational choice endogenously determines the distributions of workers' ability and of firms' technologies. Since both the firm's and the worker's problem are static, individuals choose to become a worker if $w(\bar{M})a > \pi(M(a, \bar{M}))$. Given the wage rate and the state of aggregate technology, the known value of an agent's ability thus is sufficient for the choice. A population ability distribution then implies, via labor market clearing, an occupational choice for each a and corresponding distributions of workers' ability and firms' productivity.

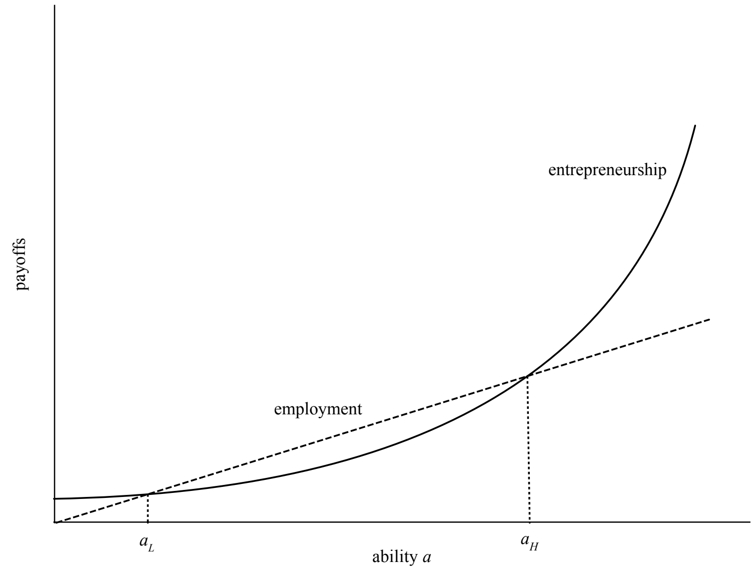
Because profits are continuous, increasing and convex in a, while wages are linear in a, it is clear that there is a threshold a_H above which it is optimal to become an entrepreneur. If $a_H < \bar{a}$ (the upper bound on a), high-productivity firms are active in the economy. At the same time, from (2) and the assumptions on M, it follows that $\pi(M(0, \bar{M})) > 0 = w \cdot 0$, so that agents with ability between 0 and a threshold a_L become entrepreneurs. Individuals with $a \in (a_L, a_H)$ choose to become workers.

The resulting occupational choice pattern is depicted in Figure 7, which plots the value of entrepreneurship (solid line) and of employment (dashed line) against a. Agents with a above a_H or below a_L become entrepreneurs, and individuals with intermediate a choose to become workers. When there is additional heterogeneity that is orthogonal to that in a, e.g. differences in taste for entrepreneurship as in Section 5, this pattern persists in the sense that the probability that entrepreneurship is the optimal choice is higher for high and low levels of a than for intermediate levels.

This two-sided occupational choice pattern differs markedly from the pattern usually obtained in models in the spirit of Lucas (1978), where only the individuals with the highest entrepreneurial ability choose entrepreneurship. The self-employed in Gollin (2007) also have relatively high entrepreneurial ability and potential wages.

Yet empirical evidence clearly suggests that entrepreneurs tend to be drawn from both extremes of the ability distribution. For instance, Gindling and Newhouse (2012) show, using

Figure 7: The payoffs to employment and entrepreneurship



household data from 98 countries, that average education, household income and consumption are highest among employers and lowest among own-account workers, with employees lying in between. Poschke (2013) finds a similar pattern in U.S. National Longitudinal Survey of Youth (NLSY) data. In addition, the firm size distribution in any country is dominated by small firms. As a consequence, a model built to address the firm size distribution needs to capture the empirical selection pattern, which includes low-ability entrepreneurs.[23]

Equilibrium. An equilibrium of this economy consists of a wage rate w and an allocation of agents to activities such that, taking w as given, agents choose optimally between employment and entrepreneurship, firms demand labor optimally, and the labor market clears.

Denoting the density of firms over a by $\nu(a)$, their total measure by B, total effective labor supply by $N \equiv \int_{a_L}^{a_H} a\phi(a)\mathrm{d}a$, and defining $\eta = \frac{1}{\sigma-1}\frac{\gamma}{1-\gamma}$, the equilibrium wage rate then is obtained from labor market clearing as

$$w(\bar{M}) = \gamma \left[\frac{B}{N}\int \nu(a)M(a,\bar{M})^{\eta}\mathrm{d}a\right]^{1-\gamma}, \tag{LM}$$

where here and in the following the integral is over the set of entrepreneurs, $a \in [0, a_L] \cup$

[23]In the model, the activity of low-profit entrepreneurs is due to the specific way in which technology and its relationship with ability is modelled here. Yet, while the specification chosen here delivers the existence of low-productivity firms somewhat directly, their owners' occupational choice arises naturally in more general settings with heterogeneity in productivity and pre-entry uncertainty about a project's merits, as shown in Poschke (2013). (Astebro, Chen and Thompson (2011) and Ohyama (2012) also present theories predicting that entrepreneurs are more likely to come from the extremes of the ability distribution.)

$[a_H, \bar{a}]$. The free entry or optimal occupational choice condition $w(\bar{M})a_i = \pi(a_i, \bar{M}), i = L, H$ can be rewritten as

$$w = \left[(1-\gamma)\frac{M(a_L, \bar{M})^\eta}{a_L}\right]^{1-\gamma}\gamma^\gamma = \left[(1-\gamma)\frac{M(a_H, \bar{M})^\eta}{a_H}\right]^{1-\gamma}\gamma^\gamma. \tag{FEC}$$

Equilibrium can be represented as the intersection of (LM) and (FEC) in a_L, w-space. a_H then follows from the second equality in (FEC). Since (LM) implies a strictly positive relationship between a_L and w and (FEC) a strictly negative one, a unique equilibrium exists for any $\bar{M}$.

It is also useful to combine (LM) and (FEC), which yields

$$\begin{aligned} \frac{B}{N}\int \nu(a)M(a, \bar{M})^\eta \mathrm{d}a &= \frac{1-\gamma}{\gamma}\frac{M(a_L, \bar{M})^\eta}{a_L} \\ &= \frac{1-\gamma}{\gamma}\frac{M(a_H, \bar{M})^\eta}{a_H}, \quad a_H > a_L. \end{aligned} \tag{EQ}$$

The right hand side of these equations is convex in a (for $\bar{M} > 1$) and approaches infinity as a goes to zero or to infinity. The left hand side is finite as long as there are workers ($a_L < \min(a_H, \bar{a})$). In line with the previous paragraph, this implies that a_L and a_H are unique. While it is possible that $a_H > \bar{a}$ (the upper bound on ability in the population), any equilibrium features strictly interior a_L.

4 Development and the firm size distribution

In this model, technological improvements affect occupational choice and, through this channel, the firm size distribution. Changes in the technological frontier affect incentives to become a worker or an entrepreneur both through their effect on potential profits and on wages. As technology advances, some firms stay close to the advancing frontier, while others fall behind. As a result, profits as a function of ability change, the populations of firms and workers change, and the equilibrium wage rate changes. This section shows first the effect of technical change on occupational choice, and then on the firm size distribution.

4.1 The technological frontier and occupational choice

Equilibrium in this economy is described by (EQ). This shows that for $\bar{M} > 1$, occupational choice is characterized by two thresholds, a_L and a_H, as shown in Figure 7. In general, an increasing technological frontier $\bar{M}$ raises both sides of (EQ) for any given thresholds a_L and

a_H, as better technology raises both wages and profits. These changes affect entrepreneurs of different ability differently, given that the elasticity of profits with respect to $\bar{M}$,

$$\varepsilon(\pi(\cdot), \bar{M}) = \eta\mu(a) - \frac{\gamma}{1-\gamma}\varepsilon(w, \bar{M}), \tag{3}$$

depends on individual ability. While higher wages – the cost of inputs – affect all entrepreneurs similarly ($\varepsilon(w, \bar{M})$ denotes the elasticity of the wage rate with respect to $\bar{M}$), more able entrepreneurs receive a larger boost to their productivity from new technology, and thus see their profits increase by more. Low-ability entrepreneurs' profits decrease, as the increase in productivity does not compensate for the increase in input cost.

A technological advance makes entrepreneurship more attractive relative to employment for individuals with ability a such that $\mu(a) > (\sigma - 1)/\gamma \cdot \varepsilon(w, \bar{M})$. Using (LM), this is the case if $a > \tilde{a}$, defined by

$$\mu(\tilde{a}) \equiv \frac{\tilde{\mu}}{\tilde{M}} \equiv \int \nu(a)\mu(a)M(a, \bar{M})^\eta \mathrm{d}a \left[\int \nu(a)M(a, \bar{M})^\eta \mathrm{d}a\right]^{-1}. \tag{4}$$

For those with $a < \tilde{a}$, an increase in $\bar{M}$ makes employment more attractive relative to entrepreneurship.

The evolution of occupational choice patterns as $\bar{M}$ increases then depends on the size of a_L and a_H relative to $\tilde{a}$. Since $\tilde{a}$ could be less than a_L, lie between a_L and a_H, or exceed a_H, the dynamics of occupational choice in response to technological progress go through three stages. In a nutshell, as $\bar{M}$ increases, marginal entrepreneurs enter (exit) if their ability is high (low) relative to other active entrepreneurs. As occupational changes evolve with $\bar{M}$, these relations change.

First, in a situation where $a_H > \bar{a}$ and only low-ability entrepreneurs are active, entrepreneurs of ability a_L are the most productive ones, so that $\mu(a_L) > \mu(\tilde{a})$. Since higher $\bar{M}$ raises profits more than employment income for entrepreneurs at a_L, entrepreneurs just above a_L find it optimal to enter, and the threshold a_L rises. Clearly, entrepreneurs just below a_H also experience an increase in potential profits relative to earnings, pushing a_H down. This process continues until a_H reaches $\bar{a}$, and high-ability entrepreneurs start to be active in the economy. As long as a_L and a_H exceed $\tilde{a}$, increases in the technological frontier continue to imply higher a_L and lower a_H.

At the same time, an advancing technological frontier also raises $\tilde{a}$.[24] While $a_H > \bar{a}$, these

[24]For given thresholds a_L, a_H, the elasticity of $\mu(\tilde{a})$ with respect to $\bar{M}$ is $\eta(\tilde{M}\int \nu(a)\mu(a)^2 M(a, \bar{M})^\eta \mathrm{d}a - \tilde{\mu}^2)/(\tilde{M}\tilde{\mu})$. The term in parentheses is weakly positive by the Cauchy-Schwartz inequality, and strictly so if there is dispersion in the productivity of active firms.

increases are smaller than those in a_L. But $\tilde{a}$ increases not only because of the direct effect of technological progress, but also because of occupational choice: the continuing entry of high-ability entrepreneurs shifts the weights $\nu(a)$ to higher values of a, raising $\mu(\tilde{a})$ and thus $\tilde{a}$. (Entry of high-ability entrepreneurs raises both the numerator and the denominator of $\mu(\tilde{a})$. But the increase in the numerator dominates as long as $a_H > \tilde{a}$.) Eventually, $\tilde{a}$ reaches and then exceeds a_L.[25] At this point, the economy enters a second phase, where further increases in the technological frontier reduce profits relative to earnings for entrepreneurs at a_L. Profits continue to rise relative to earnings for entrepreneurs at a_H. Hence, both a_L and a_H decline. The set of low-ability entrepreneurs shrinks, while that of high-ability entrepreneurs grows.

Finally, as $\tilde{a}$ continues to increase and a_H continues to fall, $\tilde{a}$ eventually reaches a_H. This occurs when the entry of high-ability entrepreneurs has reduced a_H to such an extent that a_H has become a relatively low level of ability within the set of active entrepreneurs. At this point, $\tilde{a}$ continues to increase in $\bar{M}$, both due to the direct effect of $\bar{M}$ and due to declining a_L. Once $\tilde{a}$ exceeds a_H, a further improvement in technology makes it optimal for entrepreneurs with ability a_H to switch to employment, i.e. a_H rises. In this third and final phase, a_L continues to fall, and a_H rises. Both thresholds remain in the interior of the domain for ability. The following proposition summarizes the dynamics of occupational choice.

Proposition 1. *For $\bar{M} > 1$ and under the assumptions made in Section 3, the economy traverses three phases of occupational choice dynamics in sequence.*

P0 *The threshold a_L rises and a_H declines.*

P1 *Both thresholds decline.*

P2 *a_L declines and a_H rises.*

In the following, I will ignore **P0** for lack of empirical relevance.[26]

Advancing technology does not lift all boats here. By assumption, the most able agents benefit most from advances in the technological frontier, as they can deal more easily with the increased complexity and use a larger fraction of the new technologies. Low-ability entrepreneurs benefit less. In fact, increasing wages due to higher productivity at top firms

[25] A simple argument for this is by contradiction: if a_L always increased faster than $\tilde{a}$, it would keep increasing and eventually hit a_H or $\bar{a}$. This is not consistent with equilibrium.

[26] In the quantitative exercise, this phase turns out to be very short, and to occur only for values of $\bar{M}$ below those of the poorest country in the GEM sample (Uganda).

(wage earners always gain from technological improvements) mean that the least productive firms' profits fall as technology improves. As a consequence, marginal low-productivity entrepreneurs convert to become wage earners, and eventually also do better, though not necessarily immediately. The lowest-ability agents ($a = 0$) always lose. Technology improvements thus have a negative effect on low-productivity firms that operates through wage increases.

4.2 Advances in the technological frontier and the firm size distribution

Changes in occupational choice shape the evolution of the firm size distribution. The evolution of the entrepreneurship rate B is straightforward. While it rises in the empirically irrelevant phase **P0**, it is obvious that it declines in phase **P2**, since a_L declines and a_H rises in that phase. In **P1**, in which both a_L and a_H decline, B also falls. The argument is by contradiction: For B to remain unchanged, exiting low-productivity firms would have to be replaced by an equal measure of high-productivity firms. This change would also imply reduced labor supply in efficiency units. At the same time, the improvement in the productivity distribution brought about by lower a_L and a_H raises labor demand. This situation cannot be an equilibrium. In equilibrium, the exiting low-productivity firms need to be replaced by fewer high-productivity entrants, implying that B declines as $\bar{M}$ increases.

Given that the average worker supplies $N/(1-B)$ efficiency units and that the average firm uses N/B efficiency units, average employment in terms of workers is $(1-B)/B$. Since B declines with $\bar{M}$, average employment increases in $\bar{M}$, in line with both the cross-country and the time series facts.

Percentiles of the size distribution also change with $\bar{M}$ in the model. Let $\mathcal{V}(a)$ be the *cdf* associated with $\nu(a)$. The structure of occupational choice in the model implies that the decline in the mass of firms that occurs as $\bar{M}$ rises takes place in the middle of the distribution of entrepreneurial ability, as the thresholds a_L and a_H shift. As a consequence, probability mass shifts towards the extremes of the distribution. More precisely, denoting by a_i' the value of threshold a_i induced by the new, higher value of $\bar{M}$, the fraction of firms above a_L' declines as $\bar{M}$ increases. The same holds for the fraction of firms below a_H'. As a result, $\mathcal{V}(a)$ rises for any a below a_L', and falls for any a above a_H'. (In phase **P1**, these statements also hold for a_H instead of a_H'.)

The changes in $\mathcal{V}$ translate into changes in percentiles of the distribution: employment at any fixed percentile below $\mathcal{V}(a_L')$ declines, while employment at any fixed percentile above $\mathcal{V}(a_H')$ rises. This implies that the interquartile ratio unambiguously increases with $\bar{M}$ if the

75th percentile of the size distribution is above $\max(a_H, a'_H)$ and the 25th below a'_L. The model thus can match the finding of higher size dispersion in richer countries in this case.[27]

The next section goes a step further and explores the quantitative implications of the model.

5 Quantitative exercise: occupational choice and entrepreneurship across countries

How much of the variation in the firm size distribution can be attributed simply to development? To answer this question, I first calibrate the model using data for the U.S. economy, which is commonly taken to be closest to the frictionless benchmark. Given good performance in this exercise, I then evaluate how well the model, with parameters for the U.S., can account for cross-country patterns when countries differ only in aggregate technology. This gives an indication of the importance of development as a driver of differences in the firm size distribution.

5.1 Generalized model

For the quantitative exercise, it is useful to generalize the very stylized model from Section 3 slightly. I introduce two modifications: production of intermediates with capital and labor, and heterogeneity in the taste for entrepreneurship. I also describe the choice of functional form for $M(a)$.

Capital. In the simple model in Section 3, the differentiated activities used for producing final output use labor only. The aggregate input X has constant returns to scale in all labor inputs. Replace this by

$$X = \left(\int_0^{M_i} (n_j^{\alpha} k_j^{1-\alpha})^{\frac{\sigma-1}{\sigma}} \mathrm{d}j \right)^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

i.e., production of intermediates with capital and labor. This allows setting α and γ to match income shares in the data. Firms' optimization is as in Section 3, taking the wage rate w and the rental rate of capital r as given. Households, who own the capital stock and rent it to firms, now face a capital accumulation decision. Their Euler equation, evaluated at the

[27] How the standard deviation of log employment or skewness change with $\bar{M}$ is a quantitative question. While they increase monotonically in $\bar{M}$ in the quantitative exercises reported in the next section, this may depend on the choice of parameters and functional forms.

steady state of the economy they live in (thus, given its $\bar{M}$), prescribes equating the rental rate of capital net of depreciation to the rate of time preference. Assuming a common rate of time preference ρ and a common depreciation rate δ, this implies $r = \rho + \delta$. The firm's optimality condition for capital then pins down the aggregate capital stock.[28]

Taste heterogeneity. In the model of Section 3, only agents with $a < a_L$ or $a > a_H$ become entrepreneurs. Given the one-to-one mapping between a and M, this implies a bimodal firm size distribution with only low- and high-productivity firms, but no firms with intermediate productivity. This is clearly counterfactual. Incorporating heterogeneity in tastes for entrepreneurship into the model allows to "fill in" the hole in the middle of the firm size distribution, while also adding realism. Indeed, most empirical studies of entrepreneurship point to some role for heterogeneity in tastes or risk aversion for entrepreneurship (see e.g. Hamilton 2000, Hurst and Pugsley 2011).

Thus, suppose that agents differ in their taste for entrepreneurship τ. Define this such that individuals choose entrepreneurship if $\tau\pi(a) > w \cdot a$. $\tau > 1$ then implies "enjoyment" of entrepreneurship. If agents enjoy entrepreneurship, they will choose it even if $\pi(a) < w \cdot a$. Whether on average agents enjoy entrepreneurship is an empirical question; therefore the distribution of τ has to be calibrated, and the mean could be different from 1. A mean below 1 indicates that on average, individuals do not enjoy entrepreneurship.

With this additional dimension of heterogeneity, there are entrepreneurs of all levels of ability, and the productivity distribution can be unimodal if the ability distribution is so. However, individuals of high or low ability are still more likely to become entrepreneurs. Changes in $\bar{M}$ shift the relationship of $\pi(a)$ and wa and therefore the taste threshold for entering entrepreneurship, resulting in an evolution of the proportion of agents with a given a who are entrepreneurs.[29]

The technological frontier and complexity. How much additional complexity do advances in the technological frontier comport? The assumptions introduced in Section 3 imply $M(a, \bar{M}) = \kappa \bar{M}^{\mu}(a)$, where $\mu(a)$ is an increasing, weakly convex function. From here onwards, let $M(a, \bar{M})$ be

$$M(a, \bar{M}) = \bar{M}^{a+\lambda} = \bar{M}^{\lambda}\bar{M}^{a}, \quad \lambda > 0. \tag{6}$$

[28]While growth in $\bar{M}$ leads to changes in occupational choice and in the share of entrepreneurs, the setting is consistent with balanced growth since increases in $\bar{M}$ constitute labor-augmenting technical progress and the aggregate production function exhibits constant returns to scale (King, Plosser and Rebelo 1988). Results in this section can thus also be interpreted as developments along the balanced growth path of an economy.

[29]Heterogeneity in risk aversion combined with a simple extension of the model would yield similar results.

Then the elasticity of M with respect to $\bar{M}$ is $\mu(a) = a + \lambda$, which fulfills the assumptions imposed in Section 3. Given a distribution of a, the parameter λ controls the strength of the common versus heterogeneous effect of changes in aggregate technology $\bar{M}$, and thus the strength of skill-biased change in entrepreneurial technology: The higher λ, the larger the common component, and the smaller the heterogeneity in changes in M due to an increase in $\bar{M}$. Thus, if values of a are very small relative to λ, changes in M induced by changes in $\bar{M}$ are similar for all firms, and technical change exhibits little skill-bias. In the opposite case, if $\lambda = 0$, the common component is zero and heterogeneity of the effects of technical change is maximized, implying strong skill bias.

With this specification of $M(\cdot)$, the level of $\bar{M}$ given a value of λ drives aggregate output, occupational choice and in particular the dispersion of the firm size distribution. In contrast, for a given $\bar{M}$, changes in λ act like neutral technical changes and do not affect allocations. They do however affect how the firm size distribution changes as $\bar{M}$ changes. This property allows us to calibrate $\bar{M}$ and λ separately, by combining information from the cross section and the time series of the U.S. firm size distribution.

5.2 Calibration

The model is calibrated to U.S. data. Some parameters can be set using standard numbers from the literature, while the remaining ones are calibrated to match a set of moments describing the U.S. economy.

The share parameters γ and α are set to generate a profit share of income of 10% and a labor share of two thirds. This implies a γ of 0.9 and an α of 0.74. The elasticity of substitution among intermediate inputs is set to 4, which is about the 75th percentile of the distribution of σ across 4-digit industries estimated by Broda and Weinstein (2006).[30] Setting the rate of time preference to 4% and the depreciation rate to 10% per annum implies a rental rate of capital of 14%.

For the remaining parameters, first suppose that the ability and taste distributions are lognormal. A lognormal ability distribution implies that the wage distribution would be lognormal if everyone was an employee. With taste heterogeneity, entrepreneurs will come from across the ability distribution, and the wage distribution will be close to lognormal. For tastes, a lognormal distribution also seems natural, as they affect payoffs multiplicatively. Letting $\ln a \sim N(\mu_a, \sigma_a)$ and $\ln \tau \sim N(\mu_\tau, \sigma_\tau)$ and normalizing μ_a to be zero, the remaining moments to be calibrated are $\sigma_a, \mu_\tau, \sigma_\tau, \lambda$ and $\bar{M}$.[31]

[30] Results are robust to setting σ substantially higher, to 6. This is although the sensitivity of profits with respect to $\bar{M}$ declines with σ (see e.g. equation (2)).

[31] Setting $\mu_a = 0$ is a normalization because changes in μ_a can be undone by changing $\bar{M}$ and λ appropri-

Data and model moments are shown in Table 5. U.S. data is for the year 2005, or close years where data for that year is not available. To pin down the parameters, information about the firm size distribution, about the distribution of wages and about the link between the two is needed. Targets are chosen accordingly:[32] First, wage inequality, measured as the ratio between the 90th and the 10th percentile of the wage distribution, is taken from Autor, Katz and Kearney (2008, Figure 2.A) and helps to pin down σ_a. Second, average employment is informative about μ_τ, the mean taste for entrepreneurship. In an analysis of occupational choice, the broadest possible set of firms run as full-time concerns should be considered, so the target combines information from the Census Businesses Dynamics Statistics (BDS) on employer firms with CPS data on the self-employed reported in Hipple (2010) that is informative about full-time entrepreneurs without employees. Third, changes in σ_τ affect occupational choice and the firm size distribution. Concretely, increasing σ_τ implies that occupational choices are on average more taste-based. This leads to a distribution of entrepreneurs that tends to have lower ability, and therefore to a larger fraction of small firms in the economy. The share of firms with fewer than five employees thus is an informative calibration target. Fourth, as seen in the previous section, the level of $\bar{M}$ also affects the dispersion of the firm size distribution. Given the targets already chosen, this is well-captured by the share of employment in large firms (employment over 500). Both of these targets are computed by combining size class data from the Census Statistics of U.S. Businesses (SUSB) and from the CPS (for non-employers).

Finally, to separate λ and $\bar{M}$, information on changes over time is needed. Figure 6 showed several series for the evolution of average firm size, covering different subperiods. To aggregate this information, I compute the elasticity of average firm size with respect to output per worker for each series. This lies between 0.12 and 0.57. To be conservative, I target an elasticity of 0.34, which is in the middle of the range in the data. Moreover, this value is close to the ones implied by the recent BDS series (1988-2006) and by the BEA Survey of Current Business series when omitting the Great Depression years.

Values of the calibrated parameters are reported in Table 6. On average, individuals like entrepreneurship (the implied average τ in the population is clearly above 1), but there is substantial variation. The combination of these two features is required to generate the

ately.

[32]In fact, the five parameters have to be calibrated jointly. While the following discussion stresses the main informational contribution of individual targets, parameters and target choices actually interact.

Table 5: Calibration: Data and model moments

	model	data
average employment $\bar{n}$	12.0	11.9
fraction of firms with $n < 5$	0.84	0.81
fraction of employment in firms with $n \geq 500$	0.47	0.47
ln 90/10 wage ratio	1.55	1.56
$\varepsilon(\bar{n}, Y)$	0.34	0.34
not targeted:		
fraction of firms with $n \leq 1$	0.61	0.58
fraction of employer firms with...		
$n < 5$	0.59	0.55
$n < 10$	0.74	0.76
$n < 100$	0.96	0.98
entrepreneurship rate	0.08	0.08

Sources for data moments: average firm size and entrepreneurship rate from Census Business Dynamics Statistics (BDS) and CPS data as reported in Hipple (2010); size distribution from BDS tabulations and CPS data (Hipple 2010); wage ratio from Autor et al. (2008, Figure 2A); elasticity of average employment with respect to output per worker uses average firm size data plotted in Figure Figure 6 combined with data on non-farm employment from the BLS and from Weir (1992, Table D3), reprinted in Carter et al. (2006), and data on non-farm output from the BEA (`http://www.bea.gov/bea`, Table 1.3.6) and from U.S. Department of Commerce (1975, Series F128).

observed large share of very small firms. Also note that the $\bar{M}$ resulting from the calibration describes the U.S. level of technology in 2005. To evaluate cross-country patterns, it will be necessary to set other countries' $\bar{M}$ relative to the U.S. level such that the output ratios match the data.

5.3 The U.S. time series

The model does an excellent job at reproducing the U.S. firm size distribution both in 2005 and over time. The model-generated "time series" of average employment in the U.S. is plotted against non-farm output per worker in Figure 8. As the calibration fits the observed elasticity of 0.34 well, the series of average employment also fits well.[33] The lower part of Table 5 shows model and data moments of the firm size distribution in 2005 that were not targeted in the calibration. It shows that the model fits not only the two targeted related moments very closely, but also reproduces the rest of the size distribution very well.

[33] While the model fits the short series of recent data well, it is evident that it could still fit rather well if a different average size target from one of the other sources covering a narrower population of firms were used.

Table 6: Calibrated parameter values

from external sources:				
γ	α	σ	ρ	δ
0.9	0.74	4	0.04	0.1
from fitting U.S. target moments:				
σ_a	μ_τ	σ_τ	λ	$\bar{M}$
0.616	0.616	1.131	141	1.407

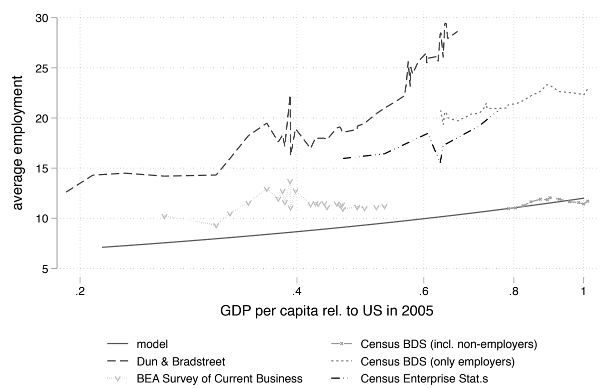


Figure 8: Average firm employment over U.S. history, data and model

Sources: Model results plus sources given in notes to Figure 6 and Table 5.

The model also does well at matching the recent evolution of some moments of the firm size and productivity distributions in the U.S.. Table 7 shows changes in employment shares and in the fraction of firms by size category for employer firms in the model, and compares them to their data counterparts from Table 4, which are reproduced here. In both model and data, there is a noticeable increase in the fraction of large firms and in their employment share.

Measures of dispersion in the model also compare well to the data. The standard deviation of log employment, computed using size bin means, increases by 0.04 in the model, compared to 0.076 in the data. The standard deviation of TFP in the model increases by 45%, compared to the number of 52% that Kehrig (2012) reports for U.S. manufacturing plants. Given that the model only targeted the change in average size, it seems fair to say that it replicates changes throughout the distribution well.[34]

[34] In each case, model statistics are computed for levels of $\bar{M}$ that generate the levels of output in the corresponding years in the data.

Table 7: Changes in the U.S. firm size distribution, 1979-2005, data and model

	firms with employment above				
Change in group ... (in percentage points)	5	50	100	250	500
employment share (data)	1.15	3.56	3.63	3.03	2.43
employment share (model)	0.88	2.38	2.80	3.16	1.95
share of firms (data)	1.63	0.67	0.39	0.15	0.07
share of firms (model)	2.56	0.86	0.50	0.20	0.15

Notes: Data sources and further remarks as in Table 4.

A frictionless model can thus account well for changes in the U.S. firm size distribution over time. Key required features are skill-biased change in entrepreneurial technology and occupational choice. Before assessing their role quantitatively, I turn to evaluating how much of the cross-country variation in the firm size distribution the model can explain purely with differences in technology.

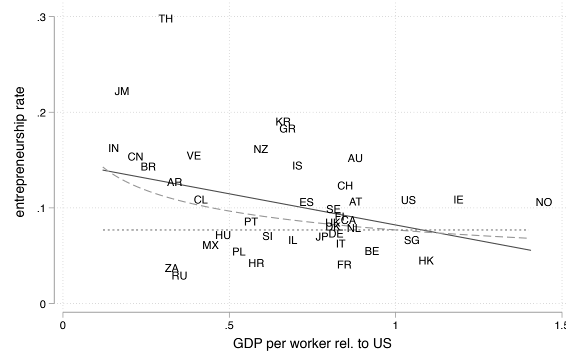
5.4 Cross-country results

To gauge the impact of development on the firm size distribution, each country is assigned the $\bar{M}$ that replicates the output per worker ratio to the U.S. observed in the data. This $\bar{M}$ is then taken to be the country's effective state of technology.[35] Figure 9 plots average firm size, the entrepreneurship rate, and firm size dispersion generated by the model for these levels of $\bar{M}$ against the data. Results for the fraction of large firms ($n \geq 10$) are shown in Figure 13 in Appendix A. The solid line in each graph is the OLS fit discussed in Section 2. The dashed lines are the outcomes generated by the benchmark model. Quantitative measures of fit are presented in Table 8. The column labelled "data" contains the regression coefficients from Table 2. The columns labelled "model" show analogous (semi-)elasticities of moments of the firm size distribution with respect to income, computed using model data.

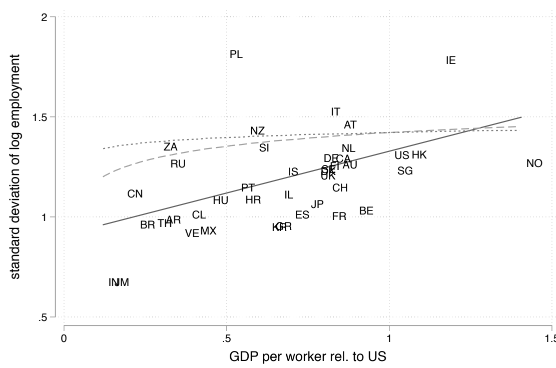
The quantitative fit to cross-country variation in the firm size distribution is good given that the model is fairly stylized and has been calibrated to the U.S.. The second column of Table 8 shows that the model can account for almost half of the elasticity of average firm size with respect to income per worker. It accounts for three quarters of the variation in the

[35]Strictly speaking, $\bar{M}$ of course also captures non-technological sources of income differences, just as TFP does. It appears reasonable that these also affect entrepreneurs' technological opportunities. Also note that cross-country differences in $\bar{M}$ are taken to be exogenous here; explaining them is beyond the scope of this paper.

(a) Average employment



(b) The entrepreneurship rate



(c) Standard deviation of log employment

Figure 9: The firm size distribution versus output per capita: data (solid) and model (benchmark: dashed, fixed occupational choice: dotted)

Notes: Data sources as in Figure 3, and measures as in Section 2. Data figures are identical to the GEM figures shown in Section 2.

entrepreneurship rate with income, and for half of the systematic variation of the standard deviation of log employment with income.[36]

Variation in moments of the firm size distribution over U.S. history suggested that such differences need not always be due to distortions. Going beyond this, the results in this section show that a frictionless model can account not only for changes in the firm size distribution over time in the U.S., but also for a good part of the cross-country variation. While some of that variation may still be due to distortions (more on that below), it is clear that not all of it need be.

[36]While the performance in terms of the elasticity with respect to income is good, it is clear from Figure 9 that the model does not match data averages for several moments. This is due to the calibration strategy, which targets U.S. statistics for average employment, and not cross-country averages or higher moments.

Table 8: Fit of the model in cross-country data: data and model elasticities with respect to GDP per capita

variable	data	model: benchmark (SCET only)	model: SCET and SDDs
log average employment	0.718	0.334	0.718
entrepreneurship rate	-0.040	-0.030	-0.099
standard deviation of log n	0.228	0.102	0.367

Notes: The table shows elasticities (first row) or semi-elasticities (remaining rows) of the indicated variables with respect to GDP per capita. Elasticities in the column labelled "data" are from the regression results using GEM data in Table 2. The benchmark column shows results for the case with skill-biased change in entrepreneurial technology (SCET) only. The column labelled "SCET and SDDs" shows results for the case with SCET and size-dependent distortions (SDDs). SDDs are set such that the country with GDP per worker relative to the U.S. matching that of China has a value of ζ of 0.3.

5.5 The importance of skill-biased change in entrepreneurial technology and occupational choice

Changes in occupational choice and the firm size distribution with $\bar{M}$ in the model are driven by a combination of the direct effect of $\bar{M}$ on relative productivity and thus the firm size distribution, and its effect on occupational choice. How important are these two components? To gauge this, I conduct two exercises. First, I compute the firm size distribution for a set of counterfactual economies, using the same levels of $\bar{M}$ as in Section 5.4, but keeping the occupational choices, and thus the measure of entrepreneurs and their distribution of ability, as in the benchmark (U.S.) economy. Second, I evaluate whether the presence of SCET affects output growth.

For the first exercise, results are shown as dotted lines in Figure 9. Of course, fixing occupational choice directly implies fixing the entrepreneurship rate and average firm size. With fixed occupational choices, all cross-country differences in other variables are driven by the direct effect of changing $\bar{M}$ on the firm productivity and size distribution. Since lower $\bar{M}$ effectively implies lower productivity dispersion, poorer economies also feature lower firm size dispersion in this scenario. However, the standard deviation of log employment changes less than half as much with output as in the benchmark. Most of the variation in firm size dispersion with $\bar{M}$ in the benchmark thus is due to changes in occupational choice.

Similar results hold when fixing occupational choices over U.S. history. When keeping occupational choices as in 1977, employment in firms with less than 5 employees declines by less than half as much as shown in Table 7. Employment in firms with more than 500 employees, in contrast, increases by more than in the model outcome shown in Table 7, and thus too much relative to the data.

Changes in the firm size distribution with development thus are only partly driven by the direct effect of skill-biased change in entrepreneurial technology on firms' relative productivities. Most of the change is due to changes in occupational choice in response to technical change.

I now turn to the second decomposition exercise. Clearly, SCET affects the evolution of the firm size distribution. How much does it affect output growth? To evaluate this, I compute what output would be for the same U.S. path of $\bar{M}$ if there were no SCET. To do so, I specify $M(a)$ as $\bar{M}^{\underline{a}+\lambda}\bar{M}_{bm}^{a-\underline{a}}$ (instead of $\bar{M}^{a+\lambda}$), where $\underline{a}$ is the average ability of entrepreneurs in the benchmark economy and $\bar{M}_{bm}$ is the value of $\bar{M}$ for the U.S. in the benchmark economy given in Table 6. Evidently, if $\bar{M} = \bar{M}_{bm}$ the two schedules for $M(a)$ coincide. For other values of $\bar{M}$, however, this specification implies that technical progress benefits all entrepreneurs equally; differences in a just imply level differences in productivity that do not change with $\bar{M}$.

Eliminating SCET in this way reduces annual output growth rate by 0.1 percentage points compared to the U.S. benchmark path. The reason is that with SCET, increased growth of productivity in high-productivity firms more than compensates for reduced growth in low-productivity ones.

Skill-biased change in entrepreneurial technology and occupational choice, the core elements of the model, are thus both crucial for the model's ability to account for observed variation in the firm size distribution. I next turn to an additional factor that can affect the firm size distribution.

6 An exploration of size-dependent distortions

The years since the publication of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) have seen a lot of empirical and quantitative work on size- or productivity-dependent distortions. Empirically, as pioneered by Hsieh and Klenow (2009), these distortions are often measured as wedges in first-order conditions in firm-level data. Their importance has typically been studied by introducing such wedges, often referred to as "taxes", in a heterogeneous-firm model and evaluating the effects on the efficiency of resource allocation and on aggregate outcomes, like productivity and output. Existing work has done this for one or a few countries at a time, or for few examples where the size and/or distribution of wedges is varied.

It is natural that size-dependent distortions (SDDs) should also affect the firm size distribution.[37] Since SDDs tend to reduce average firm size and size dispersion, they could generate patterns in line with those documented in Section 2 if they are larger in poorer countries. Of course, the variation in the U.S. firm size distribution over time shown in that section is unlikely to be due to changes in SDDs in the U.S., so the effect of SDDs on the firm size distribution across countries would plausibly be on top of SCET.

In this section, I thus explore how plausible cross-country variation in SDDs affects the firm size distribution. If SDDs are to be important for explaining cross-country income differences, I should also see their effect on the firm size distribution in the data. The simulations in this section speak to this.

For this exercise, I adopt a very simple specification of SDDs. Following Buera and Fattal-Jaef (2016), I assume that in each country, firms face revenue "taxes" τ that depend on their level of M relative to a reference level M_I:

$$1 - \tau(M) = \left(\frac{M}{M_I}\right)^{-\zeta\frac{\gamma}{\sigma-1}}. \tag{7}$$

It is straightforward to show that if firms behave optimally given τ, this implies a correlation of ζ between their log *quantity TFP* (TFPQ) and log *revenue TFP* (TFPR) as defined by Hsieh and Klenow (2009). A ζ of zero implies no distortions, whereas a ζ of 1 implies that it is optimal for all firms to choose the same size, no matter their productivity.[38]

Conveniently, estimates of ζ exist for a few countries, giving us empirical guidance on the cross-country variation of ζ. More specifically, Hsieh and Klenow (2007) document that the correlation between log TFPQ and log TFPR in China exceeds that in the U.S. by about 0.3.[39] To reflect this, I set ζ to zero for the U.S. and wealthier countries. For other countries, I assume that $\zeta = \epsilon_\zeta(\ln \bar{M} - \ln \bar{M}_{US})$, and set ϵ_ζ to 11.8, such that ζ is 0.3 in the model country with relative GDP per worker corresponding to that of China. The implied value of ζ for "India" is 0.38. I set M_I in each country such that net revenue from SDDs is zero. Results are similar for other conventions, e.g. setting M_I to keep aggregate capital unchanged as in Restuccia and Rogerson (2008). As in the previous section, all other parameters are

[37]In fact, Gourio and Roys (2014) and Garicano et al. (2017) show their effect on the French firm size distribution.

[38]Strictly speaking, ζ parameterizes productivity-dependent distortions here, and not size-dependent distortions. However, as long as $\zeta < 1$, more productivity firms continue to be larger despite facing larger distortions. Conversely, larger firms also face more distortions. Therefore I refer to size-dependent distortions, the more popular term in the literature.

[39]The correlation in the U.S. is not zero. Since there are omitted model features like investment adjustment costs (Asker, Collard-Wexler and De Loecker 2014) or size-varying markups (Melitz and Ottaviano 2008) that could induce this, I abstract from it.

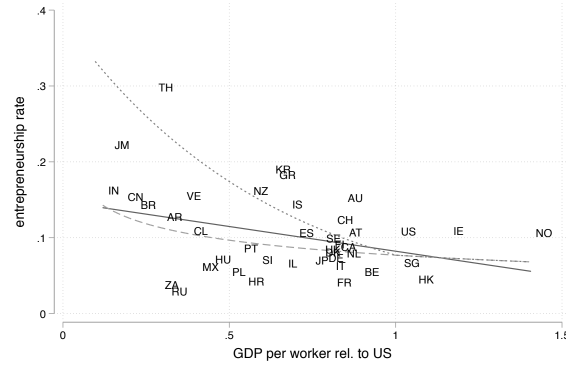
identical across countries, except for $\bar{M}$, which is set to replicate each country's GDP per worker relative to the U.S..

Results of this exercise are shown in Figure 10 (dotted lines) and the final column of Table 8. Figure 10 adds moments of the firm size distribution generated by adding SDDs to Figure 9. Consider first panel (a), plotting average firm size against income per capita. Here as in the other panels, model predictions with SDDs and without SDDs coincide for countries with the income level of the U.S. and above. For poorer countries, SDDs imply lower firm size on average, allowing the model to fit the cross-country relationship between average firm size and output per worker more closely. The last column of Table 8 shows that the model relationship mimics that in the data very closely. (Equality of the two elasticities in the first row is by coincidence.) Mean firm size in the model is still too low overall, but this is by construction as the model is calibrated to the U.S.. (This is also the reason why the entrepreneurship rate is too high in the poorest model countries. Since the predicted level of average firm size in the poorest countries is about half that in the data, the entrepreneurship rate is about double.) The fit of the share of large firms ($n \geq 10$) versus output per worker is also excellent (see Figure 13 in Appendix A).

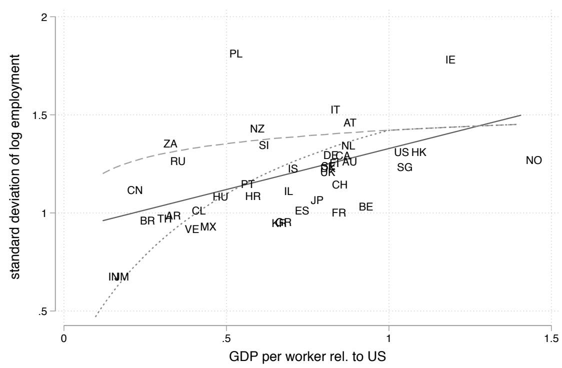
Why do these changes occur? SDDs reduce optimal employment of high-productivity firms, and raise it for low-productivity ones. Since these changes are proportional, the former changes dominate when taking the arithmetic mean, and the changes across the distribution do not simply balance in their effect on average firm size. For instance, introducing distortions with a ζ of 0.3 in the benchmark economy while keeping prices and occupational choices fixed leads to a fall in average firm size by about 50%. This is the direct effect of SDDs on the size distribution. An additional, indirect effect is via occupational choice. It arises because reduced labor demand leads to lower wages, prompting additional firm entry. Given SDDs, these entrants are mostly small. This effect leads to a decline in average firm size by roughly another 50% when introducing ζ of 0.3 in the benchmark economy, for a combined reduction of 77%. In economies with lower $\bar{M}$, the direct effect of SDDs is of comparable size, but the effect of changing occupational choices is slightly smaller due to the different starting point. For instance, introducing ζ of 0.3 in the economy with $\bar{M}$ corresponding to China, the direct effect of SDDs again is to reduce average firm size by about half, but the indirect effect only consists in a further reduction by a third.

The model including SDDs does less well in terms of dispersion. It overstates significantly how much firm size dispersion changes with income per capita. The reason for this is that in the model, as distortions become strong, size dispersion almost vanishes. Simple tweaks to how distortions vary across countries, like imposing a fixed bound on τ, do not change

(a) Average employment



(b) The entrepreneurship rate



(c) Standard deviation of log employment

Figure 10: Entrepreneurship and the firm size distribution versus output per capita: data (solid), benchmark model (dashed) and benchmark plus size-dependent distortions (dotted)

Notes: See Figure 9.

this. This tendency is inherent in SDDs. Consider the interquartile ratio as a measure of dispersion. Like average size, it declines by about half due to the direct effect of SDDs ($\zeta = 0.3$). Since SDDs favor small firms, there is a substantial shift of the distribution towards low productivity firms. As a result, both the 25th and the 75th percentile of the productivity distribution move down, but the 75th does so much more. As the productivity gap between the 75th and the 25th percentile shrinks, the interquartile employment ratio again declines by almost half. The direct effect of distortions on the firm size distribution is thus amplified by occupational choice.[40]

[40] Note that these results are not very sensitive to γ, the degree of returns to scale in the production function. At first sight, one would expect γ to be key for the effect of SDD, since optimal employment is proportional to $(1-\tau)^{\frac{1}{1-\gamma}}$. This implies that distortions are extremely powerful with γ close to 1 (be it 0.9 as here, or 0.85 as in Atkeson and Kehoe (2005)), but somewhat less so for lower values. However, the effect

Clearly, distortions affect not only the firm size distribution, but also output. By assumption, distortions are stronger in poorer countries, and therefore the effect on output there is also more pronounced. Compared to the benchmark situation, where incomes only differ because of country differences in $\bar{M}$, SDDs further reduce output per worker by 2.8% for a country with 75% of U.S. GDP per worker (similar to Japan, $\zeta = 0.055$), by 7.3% for a country at 50% (Poland, $\zeta = 0.132$), by 15% at 25% (Thailand, $\zeta = 0.277$), and by 17% at 19% (China, $\zeta = 0.3$).

Taking stock, I find that size-dependent distortions can make a substantial contribution to explaining the cross-country variation in average firm size and the importance of large firms. At the same time, their powerful effect on firm size dispersion implies that they push cross-country variation along this dimension too far.

There are two ways of reconciling these model predictions with data: it could be that the extent of SDDs is overstated in our exercise, or that there are frictions in the data that have a countervailing effect and tend to raise size dispersion in poor countries.

First, it could be that either the way SDDs are modelled here or the way they are parameterized overstate their effect. To begin, it is possible that $\zeta = 0.3$ overstates the true extent of SDDs in China. For instance, larger adjustment costs in poorer countries, for example due to more frictions in input markets, could explain part of the larger correlation between TFPQ and TFPR there, implying that the true ζ in China is smaller than 0.3. It is also possible that SDDs are smaller in the economy overall than in manufacturing, where the value of 0.3 has been estimated.

Second, other frictions that are not modelled here could in turn imply higher size dispersion in poor countries. For instance, financial frictions are more prevalent in poorer countries. They tend to impose stronger limits on the growth of small firms, thereby pushing down the size of small firms and potentially increasing size dispersion (see e.g. Beck, Demirgüç-Kunt and Maksimovic 2005, Angelini and Generale 2008). Given the potential importance of SDDs, reconciling them with observed variation in the firm size distribution across countries thus constitutes an exciting area for future work.

7 Conclusion

How and why does the firm size distribution differ across countries? This paper documents that features of the firm size distribution are strongly associated with income per capita.

of SDDs depends both on γ and on the degree of productivity dispersion in the model. Matching a given level of benchmark firm size dispersion with a lower γ requires more dispersion in the levels of productivity. This counteracts the effect of lower γ.

Firms in richer countries are on average larger, and their size is more dispersed. Firms in the U.S. are also larger, and their size more dispersed, than they were in the past.

A frictionless model of skill-biased change in entrepreneurial technology can account very well for these patterns. If more productive entrepreneurs benefit more from technological progress, development brings about a shift of economic activity to larger firms, and increases the mean and dispersion of firm size. This is driven by the direct effect of technical change, and reinforced by changing occupational choices: as better technology raises wages, marginal entrepreneurs switch to wage employment. Quantitatively, the model suggests that a substantial fraction of the cross-country variation in the firm size distribution can be attributed to differences in development.

The exploration in the last section of the paper shows that distortions, in particular size-dependent ones, may account for the remaining variation. Since there are many sources of distortions, with possibly counteracting effects on the size distribution, further investigation of their impact could be fruitful.

References

Acs, Z. J., Desai, S. and Klapper, L. (2008), 'What Does "Entrepreneurship" Data Really Show?', *Small Business Economics* **31**, 265–281.

Akyol, A. and Athreya, K. (2009), 'Self-Employment Rates and Business Size: the Roles of Occupational Choice and Credit Market Frictions', *Annals of Finance* **5**(3-4), 495–519.

Alfaro, L., Charlton, A. and Kanczuk, F. (2009), Plant-size distribution and cross-country income differences, *in* J. A. Frankel and C. Pissarides, eds, 'NBER International Seminar on Macroeconomics 2008', NBER.

Angelini, P. and Generale, A. (2008), 'On the Evolution of Firm Size Distributions on JSTOR', *American Economic Review* .

Ardagna, S. and Lusardi, A. (2009), Explaining international differences in entrepreneurship: The role of individual characteristics and regulatory constraints, *in* J. Lerner and A. Schoar, eds, 'International Differences in Entrepreneurship', University Of Chicago Press for the National Bureau of Economic Research, Chicago.

Asker, J., Collard-Wexler, A. and De Loecker, J. (2014), 'Dynamic inputs and resource (mis)allocation', *Journal of Political Economy* **122**(5), 1013–1063.

Astebro, T., Chen, J. and Thompson, P. (2011), 'Stars and Misfits: Self-Employment and Labor Market Frictions', *Management Science* **57**(11), 1999–2017.

Atack, J. (1986), 'Firm size and industrial structure in the United States during the nineteenth century', *The Journal of Economic History* **46**(02), 463–475.

Atack, J. and Bateman, F. (1999), 'Nineteenth century american industrial development through the eyes of the census of manufactures: A new resource for historical research', *Historical Methods* **32**, 177–188.

Atkeson, A. and Kehoe, P. J. (2005), 'Modeling and measuring organization capital', *Journal of Political Economy* **113**(5), 1026–1053.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Van Reenen, J. (2017), 'Concentrating on the Fall of the Labor Share', *American Economic Review Papers and Proceedings* .

Autor, D. H., Katz, L. F. and Kearney, M. S. (2008), 'Trends in US wage inequality: revising the revisionists', *Review of Economics and Statistics* **90**(2), 300–323.

Aw, B.-Y. (2002), 'Productivity dynamics of small and medium enterprises in Taiwan', *Small Business Economics* **18**(1), 69–84.

Banerjee, A. V. and Newman, A. F. (1993), 'Occupational Choice and the Process of Development', *Journal of Political Economy* **101**(2), 274–298.

Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2004), 'Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries', *World Bank Policy Research Working Paper* **3464**.

Beck, T., Demirgüç-Kunt, A. and Maksimovic, V. (2005), 'Financial and Legal Constraints to Growth: Does Firm Size Matter?', *The Journal of Finance* **60**(1), 137–177.

Bento, P. and Restuccia, D. (2017), 'Misallocation, Establishment Size, and Productivity', *American Economic Journal: Macroeconomics* **9**, 267–303.

Berlingieri, G., Blanchenay, P. and Criscuolo, C. (2017), 'The great divergence(s)', *OECD Science, Technology and Industry Policy Papers* .

Berry, A., Rodriguez, E. and Sandee, H. (2002), 'Firm and group dynamics in the small and medium enterprise sector in Indonesia', *Small Business Economics* **18**(1), 141–161.

Bloom, N., Sadun, R. and Van Reenen, J. (2012), 'Americans do IT better: US multinationals and the productivity miracle', *American Economic Review* **102**(1), 167–201.

Bollard, A., Klenow, P. J. and Li, H. (2016), 'Entry costs rise with development', *mimeo, Stanford University* .

Bonfiglioli, A., Crinó, R. and Gancia, G. (2015), 'Betting on exports: Trade and endogenous heterogeneity', *mimeo, Crei* .

Broda, C. and Weinstein, D. E. (2006), 'Globalization and the Gains from Variety', *Quarterly Journal of Economics* **121**(2), 541–585.

Buera, F. and Fattal-Jaef, R. (2016), 'The dynamics of development: Entrepreneurship, innovation, and reallocation', *manuscript* .

Buera, F. J. and Kaboski, J. P. (2012), 'Scale and the origins of structural change', *Journal of Economic Theory* **147**(2), 684 – 712.

Carter, S. B., Gartner, S. S., Haines, M. R., Olmstead, A. L., Sutch, R. and Wright, G., eds (2006), *Historical Statistics of the United States: Millennial Edition*, Cambridge University Press, Cambridge.

Chen, C. (2014), 'Essays on firm organization and international trade', *PhD thesis, Princeton University* .

Cummins, J. G. and Violante, G. L. (2002), 'Investment-Specific Technical Change in the United States (1947-2000): Measurement and Macroeconomic Consequences', *Review of Economic Dynamics* **5**(2), 243–284.

Elsby, M. W. L., Hobijn, B. and Şahin, A. (2013), 'The Decline of the U.S. Labor Share', *Brookings Papers on Economic Activity* .

Faggio, G., Salvanes, K. G. and Van Reenen, J. (2010), 'The evolution of inequality in productivity and wages: Panel data evidence', *Industrial and Corporate Change* **19**(6), 1919–1951.

Feenstra, R. C., Inklaar, R. and Timmer, M. P. (2015), 'The next generation of the penn world table', *American Economic Review* **105**(10), 3150–3182.

Felbermayr, G., Impullitti, G. and Prat, J. (2013), 'Wage Inequality, Firm Dynamics, and International Trade', *mimeo, CNRS* .

García-Santana, M. and Ramos, R. (2015), 'Distortions and the size distribution of plants: Evidence from cross-country data', *SERIES: Journal of the Spanish Economic Association* **6**, 279–312.

Garicano, L., Lelarge, C. and Van Reenen, J. (2017), 'Firm size distortions and the productivity distribution: Evidence from France', *American Economic Review* .

Garicano, L. and Rossi-Hansberg, E. (2015), 'Knowledge-based hierarchies: Using organizations to understand the economy', *Annual Review of Economics* **7**(1), 1–30.

Gindling, T. H. and Newhouse, D. (2012), 'Self-employment in the developing world', *World Bank Policy Research Working Paper* (6201).

Gollin, D. (2007), 'Nobody's business but my own: Self-employment and small enterprise in economic development', *Journal of Monetary Economics* **55**(2), 219–233.

Gourio, F. and Roys, N. (2014), 'Size dependent regulations, firm size distribution and reallocation', *Quantitative Economics* **5**(2), 377–416.

Hamilton, B. H. (2000), 'Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment', *Journal of Political Economy* **108**(3), 604–631.

Heston, A., Summers, R. and Aten, B. (2009), 'Penn World Table Version 6.3', *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania* .

Hipple, S. (2010), 'Self-employment in the United States', *Monthly Labor Review* **133**(9), 17–32.

Hsieh, C. T. and Klenow, P. J. (2007), 'Misallocation and manufacturing TFP in China and India', *SIEPR Discussion Paper* .

Hsieh, C.-T. and Klenow, P. J. (2009), 'Misallocation and Manufacturing TFP in China and India', *Quarterly Journal of Economics* **124**, 1403–1448.

Hsieh, C.-T. and Klenow, P. J. (2014), 'The Life Cycle of Plants in India and Mexico', *Quarterly Journal of Economics* **129**(3), 1035–1084.

Hurst, E. and Pugsley, B. W. (2011), 'What Do Small Businesses Do?', *Brookings Papers on Economic Activity* .

Jones, C. I. (1995), 'R&D-Based Models of Economic Growth', *Journal of Political Economy* **103**(4), 759–784.

Jovanovic, B. and Rousseau, P. L. (2008), 'Specific capital and technological variety', *Journal of Human Capital* **2**(2), 129–153.

Kehrig, M. (2012), 'Are recessions cleansing or sullying? evidence from the cyclicality of productivity dispersion', *mimeo, UT Austin* .

King, R. G., Plosser, C. I. and Rebelo, S. T. (1988), 'Production, Growth and Business Cycles: I. The Basic Neoclassical Model', *Journal of Monetary Economics* **21**(2-3), 195–232.

Klapper, L., Amit, R. and Guillén, M. F. (2010), Entrepreneurship and Firm Formation across Countries , *in* J. Lerner and A. Schoar, eds, 'International Differences in Entrepreneurship', University of Chicago Press, Chicago, pp. 129–158.

Lloyd-Ellis, H. and Bernhardt, D. (2000), 'Enterprise, Inequality and Economic Development', *Review of Economic Studies* **67**(1).

Lucas, R. E. (1978), 'On the size distribution of business firms', *Bell Journal of Economics* **9**, 508–523.

Margo, R. A. (2013), 'Economies of Scale in Nineteenth Century American Manufacturing Revisited: A Resolution of the Entrepreneurial Labor Input Problem', *NBER Working Paper* **19147**.

Melitz, M. J. and Ottaviano, G. I. P. (2008), 'Market Size, Trade, and Productivity', *Review of Economic Studies* **75**(1), 295–316.

Michaels, G. (2007), 'The division of labour, coordination, and the demand for information processing', *CEPR Discussion Paper* .

Murphy, K., Shleifer, A. and Vishny, R. (1991), 'The allocation of talent: implications for growth', *Quarterly Journal of Economics* **106**(2), 503–530.

Nugent, J. B. and Yhee, S.-J. (2002), 'Small and medium enterprises in Korea: Achievements, constraints and policy issues', *Small Business Economics* **18**(1), 85–119.

Ohyama, A. (2012), 'Entrepreneurship and Advanced Technical Knowledge', *mimeo, University of Illinois at Urbana-Champaign* .

Poschke, M. (2013), 'Who Becomes an Entrepreneur? Labor Market Prospects and Occupational Choice', *Journal of Economic Dynamics and Control* **37**, 693–710.

Restuccia, D. and Rogerson, R. (2008), 'Policy Distortions and Aggregate Productivity with Heterogeneous Establishments', *Review of Economic Dynamics* **11**(4), 707–720.

Reynolds, P., Bosma, N., Autio, E., Hunt, S., De Bono, N., Servais, I., Lopez-Garcia, P. and Chin, N. (2005), 'Gobal Entrepreneurship Monitor: Data Collection Design and Implementation 1998-2003', *Small Business Economics* **24**, 205–231.

Romer, P. M. (1987), 'Growth based on increasing returns due to specialization', *American Economic Review* **77**(2), 56–62.

Roys, N. and Seshadri, A. (2013), 'Economic development and the organization of production', *mimeo, University of Wisconsin, Madison* .

Scopp, T. S. (2003), 'The Relationship Between the 1990 Census and Census 2000 Industry and Occupation Classification Systems', *U.S. Census Bureau Technical Paper* **65**.

Summers, R. and Heston, A. (1991), 'The Penn World Table (Mark 5): An Expanded Set of International Comparisons: 1950-1988', *Quarterly Journal of Economics* **56**, 327–368.

Tomlin, B. and Fung, L. (2012), 'Exchange Rate Movements and Productivity Distribution', *Bank of Canada Working Paper* **2010-25**.

Tybout, J. R. (2000), 'Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?', *Journal of Economic Literature* **38**(1), 11–44.

Urata, S. and Kawai, H. (2002), 'Technological progress by small and medium enterprises in Japan', *Small Business Economics* **18**(1), 53–67.

U.S. Department of Commerce (1975), *Historical Statistics of the United States: Colonial Times to 1970*, U.S. Government Printing Office, Washington, DC.

Weir, D. R. (1992), A Century of U.S. Unemployment, 1890-1990, *in* R. L. Ransom, R. Sutch and S. B. Carter, eds, 'Research in Economic History', JAI Press, Greenwich, pp. 341–343.

Wiboonchutikula, P. (2002), 'Small and medium enterprises in Thailand: Recent trends', *Small Business Economics* **18**(1), 211–224.

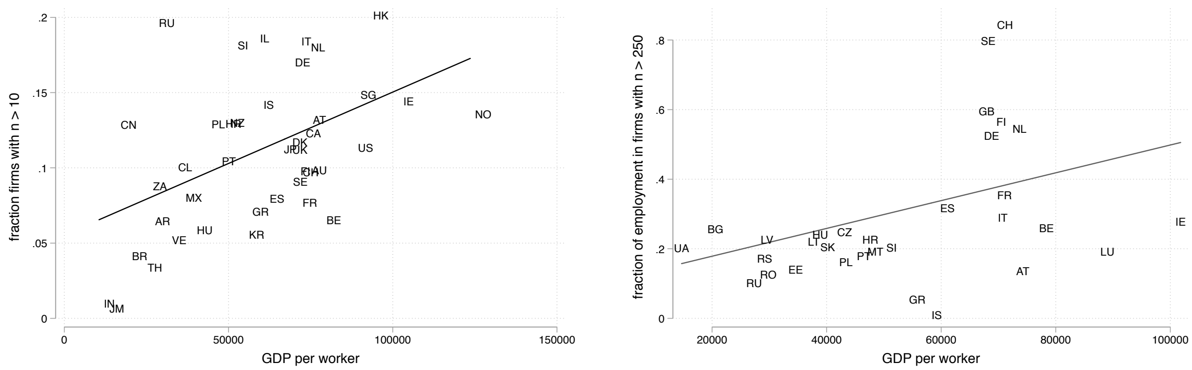
Appendix

A Additional Tables and Figures

Table 9: Country codes

country code	country name	in GEM	in Amadeus	country code	country name	in GEM	in Amadeus
AR	Argentina	X		JM	Jamaica	X	
AT	Austria	X	X	JP	Japan	X	
AU	Australia	X		KR	South Korea	X	
BA	Bosnia & Herzegovina		X	LT	Lithuania		X
BE	Belgium	X	X	LU	Luxembourg		X
BG	Bulgaria		X	LV	Latvia	X	X
BR	Brazil	X		MT	Malta		X
BY	Belorussia		X	MX	Mexico	X	
CA	Canada	X		NL	Netherlands	X	X
CH	Switzerland	X	X	NO	Norway	X	
CL	Chile	X		NZ	New Zealand	X	
CN	China	X		PL	Poland	X	X
CZ	Czech Republic		X	PT	Portugal	X	X
DE	Germany	X	X	RS	Serbia		X
DK	Denmark	X		RU	Russia		X
EE	Estonia		X	RU	Romania	X	X
ES	Spain	X	X	SE	Sweden	X	X
FI	Finland	X	X	SG	Singapore	X	
FR	France	X	X	SI	Slovenia	X	X
GR	Greece	X	X	SK	Slovakia		X
HK	Hong Kong	X		TH	Thailand	X	
HR	Croatia	X	X	TW	Taiwan	X	
HU	Hungary	X	X	UA	Ukraine		X
IE	Ireland	X	X	UG	Uganda	X	
IL	Israel	X		UK	United Kingdom	X	X
IN	India	X		US	United States	X	
IS	Iceland	X	X	VE	Venezuela	X	
IT	Italy	X	X	ZA	South Africa	X	

Note: Bosnia & Herzegovina, Belorussia, Taiwan and Uganda are excluded from the analysis in Section 2.2 due to lack of data on agricultural value added and employment.

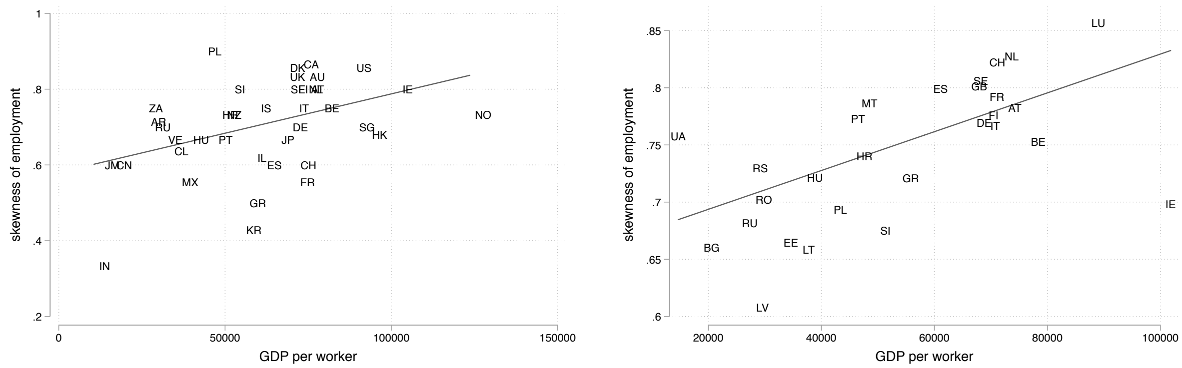


(a) The fraction of firms with $n > 10$ (GEM)

(b) The fraction of employment in large firms ($n \geq 250$) (Amadeus)

Figure 11: The importance of large firms and income per worker.

Notes: Data sources as in Figure 3. Panel (b) shows total employment in firms with at least 250 employees in Amadeus over private sector employment, computed as total employment minus general government employment from the World Development Indicators. (Results are very similar when not adjusting for government employment.) The lines represent the linear best fits. Regression results for each moment and the log of GDP per worker are reported in Table 10.



(a) Small and medium sized firms (GEM, $n < 250$)

(b) Large firms (Amadeus, $n \geq 250$)

Figure 12: Skewness and income per worker.

Notes: Data sources and further remarks as in Figure 3. Denoting the x^{th} percentile of the firm size distribution by p_x, the 90/10 percentile skewness measure used here is $((p_{90} - p_{50}) - (p_{50} - p_{10}))/(p_{90} - p_{10})$. Panel (a) excludes Brazil and Thailand, where $p_{50} = p_{10}$, implying skewness of 1. Panel (b) excludes the Czech and Slovak Republics, where $p_{50} = p_{10}$, implying skewness of 1, and excludes firm with $n < 250$. The pattern is sensitive to the inclusion of very small firms, but qualitatively similar when using data above a cutoff of 50 or higher. Regression results for each moment and the log of GDP per worker are reported in Table 10.

Table 10: The firm size distribution and income per worker – additional moments.

	GEM data			Amadeus data		
Moment	Coeff.	SE	R^2	Coeff.	SE	R^2
Fraction of firms with $n > 10$	0.046***	(0.012)	0.292			
Fraction of employment in firms with $n \geq 250$				0.182**	(0.076)	0.166
skewness of employment	0.107***	(0.032)	0.235	0.077***	(0.020)	0.350

Notes: Data sources as in Figure 3. The table shows coefficients from bivariate regressions of each moment on log GDP per capita, the standard errors on those coefficients, and the R^2 for each regression. A constant is also included in each regression (coefficient not reported). The preceding figures show these relationships for the level instead of the log of GDP. Skewness is measured as the 90/10 percentile skewness, defined as $((p_{90} - p_{50}) - (p_{50} - p_{10}))/(p_{90} - p_{10})$, where p_i stands for the i^{th} percentile of the employment distribution. Skewness results using Amadeus data are for firms with $n \geq 250$. The coefficient is statistically significantly positive at the 5% level for cutoffs from $n \geq 50$ upwards. *** (**) [*] denotes statistical significance at the 1% (5%) [10%] level.

Table 11: The firm size distribution and income per worker – using U.S. sector weights.

	GEM data			Amadeus data		
Moment	Coeff.	SE	R^2	Coeff.	SE	R^2
Entrepreneurship rate	-0.037*	(0.019)	0.110			
Log average employment	0.720***	(0.172)	0.360	0.951***	(0.307)	0.256
Standard deviation of log employment	0.269***	(0.050)	0.483	0.220**	0.087	0.189

Notes: Data sources as in Figure 3, and remarks as in Table 2.

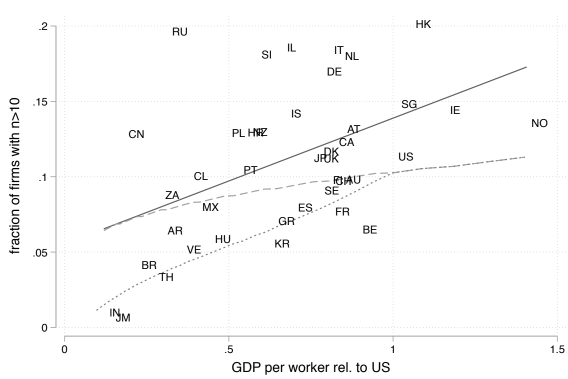


Figure 13: The importance of large ($n \geq 10$) firms and income per worker: data (solid), benchmark model (dashed), and benchmark plus size-dependent distortions (dotted)

Notes: Data sources as in Figure 3, and remarks as in Table 2.