

University of Toronto
Department of Economics



Working Paper 642

On Average Establishment Size across Sectors and Countries

By Pedro Bento and Diego Restuccia

August 05, 2019

On Average Establishment Size across Sectors and Countries[†]

Pedro Bento
Texas A&M University*

Diego Restuccia
University of Toronto
and NBER**

August 2019

ABSTRACT

We construct a new dataset for the average employment size of establishments across sectors and countries from hundreds of sources. Establishments are larger in manufacturing than in services, and in each sector they are larger in richer countries. The cross-country income elasticity of establishment size is remarkably similar across sectors, about 0.3. We discuss these facts in light of several prominent theories of development such as entry costs and misallocation. We then quantify the sectoral and aggregate impact of entry costs and misallocation in an otherwise standard two-sector model with endogenous firm entry, firm-level productivity, and sectoral employment shares. We find that observed measures of misallocation account for the entire range of establishment-size differences across sectors and countries and almost 50 percent of the difference in non-agricultural GDP per capita between rich and poor countries.

Keywords: establishment size, manufacturing, services, distortions, misallocation, productivity.
JEL codes: O1, O4, O5, E02, E1.

[†]For helpful comments we thank the editor, two anonymous referees, Alvaro Garcia Marin, and seminar participants at the National University of Singapore and the University of Texas at Austin, as well as participants at the Workshop for Productivity Dynamics, Firm Growth and Factor Reallocation at the Banco Central de Chile, the CEPR-EIEF Conference on Firm Dynamics and Economic Growth at the Banco d'Italia, and at the RIDGE Workshop on Macroeconomics and Development in Buenos Aires. All errors are our own. Restuccia gratefully acknowledges the support from the Social Sciences and Humanities Research Council of Canada and the Canada Research Chairs program.

*Department of Economics, Texas A&M University, 3056 Allen Building, 4228 TAMU, College Station, TX 77843, USA. E-mail: pbento@tamu.edu.

**Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada. E-mail: diego.restuccia@utoronto.ca.

1 Introduction

We construct a newly-assembled dataset for the average employment size of service-sector establishments for up to nine service industries across a large sample of countries. We use census or representative survey data from hundreds of sources. We combine this data for the service sector with that of the manufacturing sector from [Bento and Restuccia \(2017\)](#) to provide a more comprehensive view of establishment size in the non-agricultural sector across countries. We show that average establishment size is generally larger in manufacturing than in services, and in each sector is strongly positively related with the level of development. In particular, the cross-country income elasticity of establishment size is positive and remarkably similar across sectors of around 0.3. A critical element in the construction of our dataset for international comparisons is the inclusion of *all* establishments regardless of whether they are registered or informal and have paid employees or are self-account businesses as there is systematic variation in these categorizations across countries that can bias the relationship between establishment size and development.

Average establishment size is an important outcome in several prevailing theories of development and aggregate productivity. An important contribution of our paper is constructing a comprehensive dataset of average establishment size in the service sector across countries. Together with similarly constructed data for manufacturing, our unique dataset provides internationally comparable size data across a large set of countries for the non-agricultural sector. This is relevant as there is a growing volume of research looking beyond the manufacturing sector by analyzing also the service sector, which is reflective of the fact that the non-agricultural sector—and the service sector in particular—is becoming the dominant economic activity as a result of the process of structural transformation in developed and developing countries. Our dataset can prove useful in calibrating quantitative models, in assessing alternative theories, and guiding research towards the most relevant institutional factors determining establishment size and other aggregate outcomes. Quantitative analyses of establishment size have thus far focused on non-comparable data across countries, datasets that typically restrict the minimum

size of establishments, rendering biased comparisons of most firms in rich countries with few unrepresentative large firms in poor and developing countries.

We contribute to this literature by as a first step documenting the empirical cross-country relationships between our measures of average establishment size and several country- and sector-specific variables, such as measures of sectoral shares, openness to trade, external finance, firing costs, and firm-level distortions. In choosing potential determinants of average size, we focus on variables that are relevant to widely-used models. In many cases, competing models have different qualitative implications for average size, whereas in other cases, the implications for average size are unambiguous in the theory, both within and across sectors. The empirical finding that is most closely aligned with current theory is the negative relationship between average establishment size and the extent of misallocation across countries in each sector. The measure of misallocation we focus on is the productivity elasticity of distortions across establishments—correlated distortions for short, which we calculate using micro data from the World Bank’s Enterprise Surveys for a large number of countries. Models with endogenous entry and investment in establishment-level productivity are most closely aligned with the evidence as in these frameworks correlated distortions reduce aggregate productivity by misallocating factors across heterogeneous establishments, reducing investment in productivity, and reducing average establishment size (Hsieh and Klenow, 2014; Bento and Restuccia, 2017). Importantly, we highlight empirical cross-country relationships between average establishment size and potential determinants that appear at odds with current theories. For example, in widely-used quantitative models with exogenous productivity, financial frictions and higher firing costs are predicted to increase average size, while our data suggests the opposite. But to the extent that financial frictions and firing costs discourage productivity investment by more productive establishments, incorporating endogenous productivity into these models may generate outcomes that are more aligned with the empirical relationships we document.

In light of this empirical evidence and because our data indicates that misallocation around the world is characterized by larger productivity elasticities of distortions in both sectors in

poorer countries, we develop a two-sector model of non-agriculture featuring heterogeneous establishments making entry and productivity decisions. We focus on productivity investment by entering establishments, abstracting from investment after entry. [Hsieh and Klenow \(2014\)](#) and [Bento and Restuccia \(2017\)](#) find that incorporating life-cycle productivity investment does little to amplify the effect of policy distortions on aggregate productivity, beyond the impact of factor misallocation. Our model builds on the frameworks of [Hopenhayn \(1992\)](#) and [Restuccia and Rogerson \(2008\)](#). Upon entering the market, establishments pay a cost to increase their productivity. After this investment, an idiosyncratic productivity is realized. Establishments face policy distortions that are related to their overall productivity and hence entering establishments consider these distortions when making productivity investments. Consumers have preferences over consumption of manufacturing goods and services, and relative productivities are key determinants of sectoral allocations. Consistent with evidence, average establishment size in the model does not depend on cross-country secular differences in aggregate productivity or sectoral employment shares.

We calibrate a benchmark economy to U.S. data and study how variations in the productivity elasticity of distortions affect sectoral establishment size and productivity as well as aggregate outcomes. Our analysis shows that empirically-reasonable variations in the productivity elasticity of distortions across rich and poor countries in each sector generate substantial differences in average establishment size and productivity in each sector as well as differences in employment shares and aggregate output across economies. For instance, when we increase the productivity elasticity of distortions in each sector from 0.09 in the U.S. benchmark economy to 0.7 (a level observed for many developing countries in our data), average establishment size drops from 22 to 8 persons engaged in the manufacturing sector and from 4.8 to 1.7 persons engaged in the service sector, a reduction in size that is consistent with and close in magnitude to the evidence we document across countries. The increase in distortions also reduces sector productivity by 57 percent, only half of which is driven by factor misallocation. Because the impact of misallocation is slightly larger in manufacturing than in services, due to larger productivity

dispersion in manufacturing, sectoral output falls by more in manufacturing (54 percent) than in services (47 percent) compared to the benchmark economy. We find that accounting for entry investment roughly doubles the impact of correlated distortions on aggregate output, relative to the impact of factor misallocation in a static environment. Overall, our results suggest an important link between policy distortions (misallocation) and technology differences across countries that substantially contribute to large differences in output per capita ([Restuccia and Rogerson, 2017](#)).

Our paper builds from the seminal works of [Lucas \(1978\)](#) and [Hopenhayn \(1992\)](#). As such, our paper is related to a quantitative literature analyzing the size of establishments in development ([Barseghyan and DiCecio, 2011](#); [Moscoso Boedo and Mukoyama, 2012](#)). A related literature emphasizes the size of the informal sector arising from financial frictions, taxes, and regulations that generate misallocation ([D’Erasmus and Moscoso Boedo, 2012](#); [Leal Ordóñez, 2014](#)). [Hopenhayn \(2016\)](#) provides a systematic evaluation of variants of the Lucas and Hopenhayn models for average establishment size and [Tybout \(2000\)](#) provides a broader survey of theories of establishment size in the manufacturing sector. Our analysis is closely related to the literature on misallocation and aggregate productivity, and within this literature to the work emphasizing the dynamic implications of misallocation such as [Hsieh and Klenow \(2014\)](#), [Bento and Restuccia \(2017\)](#), and [Guner et al. \(2018\)](#). These papers focus on correlated distortions as we do, rather than the total dispersion in distortions as in most of the literature. But the literature has thus far focused primarily on the manufacturing sector due to data availability. An important departure of our paper is that we emphasize the sectoral and aggregate implications of misallocation by documenting correlated distortions and average establishment size in both manufacturing and services for a large set of countries. Our analysis provides a systematic quantitative evaluation of misallocation as a driver of differences in establishment size and productivity across sectors and countries. [Dias et al. \(2016\)](#) study factor misallocation in manufacturing and services focusing on Portugal, whereas our analysis includes a large set of countries. More importantly, we document and analyze establishment size differences across

countries and sectors. [Kumar et al. \(2001\)](#) also analyze the empirical determinants of average establishment size across countries but do so in a sample that does not include all establishments and that comprises only 15 relatively developed countries in Europe.

The rest of the paper proceeds as follows. In the next section we document a newly constructed dataset on average establishment sizes in manufacturing and services sectors across a large set of countries. [Section 3](#) discusses the main determinants of average establishment size differences across sectors and countries and relates these facts to prominent theories of development. In [Section 4](#) we describe a two-sector model that features endogenous sectoral employment and production heterogeneity in each sector. The model also features endogenous entry and productivity investment by establishments. [Section 5](#) calibrates a benchmark economy to U.S. data in order to assess the quantitative role of distortions on establishment size and productivity in each sector, as well as sectoral employment shares and aggregate outcomes. We conclude in [Section 6](#).

2 Average Establishment Size

We describe the construction of a newly-assembled dataset for the average employment size of service sector establishments across a large sample of countries using census or representative survey data, and show that average establishment size is strongly positively related to the level of development. We combine this dataset with data for the manufacturing sector from [Bento and Restuccia \(2017\)](#) to provide a more comprehensive view of establishment sizes across sectors and countries.

2.1 Data

We construct a dataset of the average employment size of service sector establishments across countries using hundreds of reports from economic censuses and nationally-representative surveys. We provide a list of countries included with the sources used for each country in Appendix

A. Our goal in the construction of this dataset is to obtain an internationally-comparable measure of average establishment size for a large sample of countries that is representative of the world income distribution, for both manufacturing and services. The primary challenges are: data availability, which typically biases country samples towards rich countries; and international comparability, due to data reported using different definitions of employment and production units, data that disproportionately include larger firms, or especially in the case of services, having data aggregated across different groups of industries. The four measures of workers and production units we use are employees, persons engaged, establishments, and firms. Employees are workers who are formally paid by a business, while persons engaged is a broader measure of employment, including unpaid (sometimes family) workers and owner-managers. In the aggregate, persons engaged accounts for all persons working in a sector. Establishments include businesses with a fixed location, as well as businesses operating out of households when a sign is posted on the premises.¹ Firms are groups of (at least one) establishment with common ownership and control.

To assess the relationship between establishment size and development, it is crucially important that the data include all establishments regardless of whether the establishments are registered or not, and whether the establishments have paid employees or not, as a substantial portion of establishments in poor countries are unregistered and own account businesses and may include unpaid family workers. For example, paid employees account for only one quarter of persons engaged in Yemen's retail sector, while paid employees account for 95 percent of workers in U.K. retail establishments. Similarly, in Sierra Leone, 83 percent of manufacturing establishments have no paid employees, and in Ghana, unpaid workers account for almost half of the manufacturing workforce. In rich countries, by contrast, paid employees account for the

¹The United States is an exception since nonemployer statistics do not distinguish between nonemployers with and without physical establishments. For example, a self-employed street vendor without paid employees would be counted as a retail establishment. Likewise, a full-time visual artist working from home (but not selling from home) would be counted as an art, culture, and recreation establishment. To address this issue in the U.S. data, we use the U.S. Bureau of Labor Statistics' Occupational Employment Statistics for 2007 to identify occupations that are likely to be counted as nonemployers, and remove these from our establishment counts. By industry, we remove door-to-door sales workers and street vendors from Retail; taxi drivers from Transportation and Storage; real estate agents and photographers from Real Estate and Business Services; artists, musicians, and authors from Art, Culture, and Recreation; and personal financial advisors from Finance and Insurance.

bulk of persons engaged. As a result of these differences between establishments in rich and poor countries, excluding nonemployer establishments generates a highly distorted picture of establishment size differences across countries. Throughout our data collection process, we have made an effort to search for evidence from methodology documents and other published reports that small establishments are not included. Any country for which such evidence exists is not included in our sample. For our manufacturing dataset, we include all countries with publicly-available data representative of all manufacturing establishments or firms.² For services, we collect data for as many as nine service industries per country: retail; wholesale; auto repair and sales; accommodations and restaurants; finance and insurance; transportation and storage; information and communications; art, culture, and recreation; and real estate and business services. We exclude education, health care, and public administration from all countries due to wildly different levels of government control and involvement. We also exclude construction to avoid ambiguity about the definition of establishment size, given the ubiquitousness of sub-contractors. We include all countries with data representative of all establishments in at least one service industry and collect data for the year closest to 2007. For the manufacturing sector we had collected data only for the entire sector for a number of years, however, given the heterogeneity of service industries we focus instead on collecting data for multiple industries within the service sector.

Our standardized definition of size is the average number of persons engaged per establishment across all establishments. For many of the countries in our service sector sample, the data are reported as total number of persons engaged and total number of establishments. But for some industries in some countries the data are reported differently as the total number of employees, the total number of firms, or a combination of these instead of persons engaged and establishments. Table 1 summarizes the sample of countries and the reported data.³

²We also include in the dataset all territories such as French Guiana, Hong Kong, and Puerto Rico. We use the word “country” solely for ease of exposition.

³GDP per capita (adjusted for purchasing power parity, PPP) for most countries is from Penn World Table v. 8.0, the IMF’s World Economic Outlook 2013, and the CIA World Factbook. For overseas departments of France, GDP per capita is from France’s National Institute of Statistics and Economic Studies and is made relative to the U.S. GDP per capita using market exchange rates. GDP per capita for Åland Islands is from Statistics and Research Åland, and adjusted for purchasing power parity using Finland’s PPP exchange rate

Table 1: SAMPLE OF COUNTRIES FOR SERVICES

Variable	Total Number of Countries	Number of Poor Countries	Number of Rich Countries
persons engaged	97	51	46
employees	78	30	48
engaged and employees	48	17	31
establishments	60	32	28
firms	80	37	43
establishments and firms	13	5	8

Notes: ‘Poor’ and ‘Rich’ refer to countries with GDP per capita below and above the median. ‘Number of Countries’ refers to number of countries in which at least one industry reports the variable in question. Data from multiple sources, see text for details.

To standardize the measure of size in each industry, we impute persons engaged per establishment using the reported data as follows. To impute the number of persons engaged in countries that only report paid employees in a particular service industry, we regress persons engaged on employees using all countries for which both measures are reported for that industry. We then use the resulting coefficient to calculate the number of persons engaged for each country-industry that reports only employees. For example, for the wholesale sector, we regress (logged) total persons engaged on (logged) employment using 36 countries that report both measures of workers, obtaining a coefficient estimate of 1.16. For countries that only report employment in the wholesale sector, we then impute (logged) persons engaged as (logged) employment times 1.16. To impute the number of establishments in an industry for which only the number of firms is reported, we follow an analogous procedure.

Our measure of average establishment size for the entire service sector is persons engaged per establishment across all service industries. One issue that arises in constructing this size measure for the entire service sector is that many countries report data for some but not all service industries. We therefore take the following steps to construct a comparable measure of establishment size in services across countries. First, we calculate average establishment size

from Penn World Table v8.0.

in a country across all service industries for which we have data. Second, we calculate average size across the same group of industries in the United States. Third, we take the ratio of average size in the country to the corresponding number in the United States and multiply this ratio by the average size in the United States across *all* service industries (equal to 5 persons engaged per establishment). This renders a comparable measure of average establishment size in the service sector across countries even if countries have data for only a subset of industries. This adjustment is important as there are substantial differences in average establishment size across industries within the service sector. For instance, Table 2 reports the average size of U.S. establishments in each service industry, which vary from 2 persons engaged in Art, Culture, and Recreation to 15 persons engaged in Accommodations and Restaurants.

Table 2: AVERAGE ESTABLISHMENT SIZE BY SERVICE INDUSTRY, UNITED STATES

Wholesale	9
Retail	6
Automobile Related	8
Accommodations and Restaurants	15
Transportation and Storage	5
Information and Communication	10
Real Estate and Business Services	3
Art, Culture, and Recreation	2
Finance and Insurance	5

In our final dataset we report the average of persons engaged per establishment for each country in services, resulting in a final sample of 127 countries. For manufacturing, we report the same measure of establishment size for a final sample of 134 countries.

2.2 Establishment Size and Development

Table 3 reports some descriptive statistics concerning average establishment size from our dataset and GDP per capita. Establishment sizes differ substantially across countries both in the manufacturing and services sectors. In addition, while average establishment sizes in the

Table 3: DESCRIPTIVE STATISTICS

	Mean	Median	Poorest Quartile	Richest Quartile
Establishment Sizes (persons engaged)				
Wholesale	6	5	7	8
Retail	4	3	2	6
Automobile Related	5	4	4	6
Accommodations and Restaurants	8	4	4	10
Transportation and Storage	9	5	6	16
Information and Communication	15	8	5	43
Real Estate and Business Services	5	4	4	6
Art, Culture, and Recreation	5	3	2	10
Finance and Insurance	14	8	9	20
Services	4	3	2	6
Manufacturing	12	9	6	17
GDP per capita (thousands)	23	17	3	52

Notes: “Poorest” and “Richest” quartiles refer to the quartiles of countries with the lowest and highest GDP per capita. Data from multiple sources, see text for details.

broader service sector are generally lower than in manufacturing—about one third of the size on average—establishment size in each sector is systematically lower in poor compared to rich countries, also a factor difference of 3. For example, in the poorest quartile of countries in our sample, average establishment size is 2 persons engaged in services compared to 6 persons engaged in the richest quartile of countries. In manufacturing, the difference is 6 persons engaged in poor countries compared to 17 persons engaged in rich countries. These patterns hold for individual service industries, with some industries featuring very large differences in average establishment size between the poorest and richest countries. For instance, the sub-industry of Information and Communication stands out with a factor difference in establishment size of almost 9.

Figure 1 documents the average establishment size in manufacturing and in services for every country in our sample against GDP per capita. The data clearly show a positive relationship between average establishment size and GDP per capita in both sectors. The elasticity of

establishment size with respect to GDP per capita is 0.30 in manufacturing and 0.32 in services. The regression slope coefficient (standard error) in Figure 1A is 0.30 (0.05) and in Figure 1B is 0.32 (0.04). The corresponding numbers of countries included are 134 and 127.

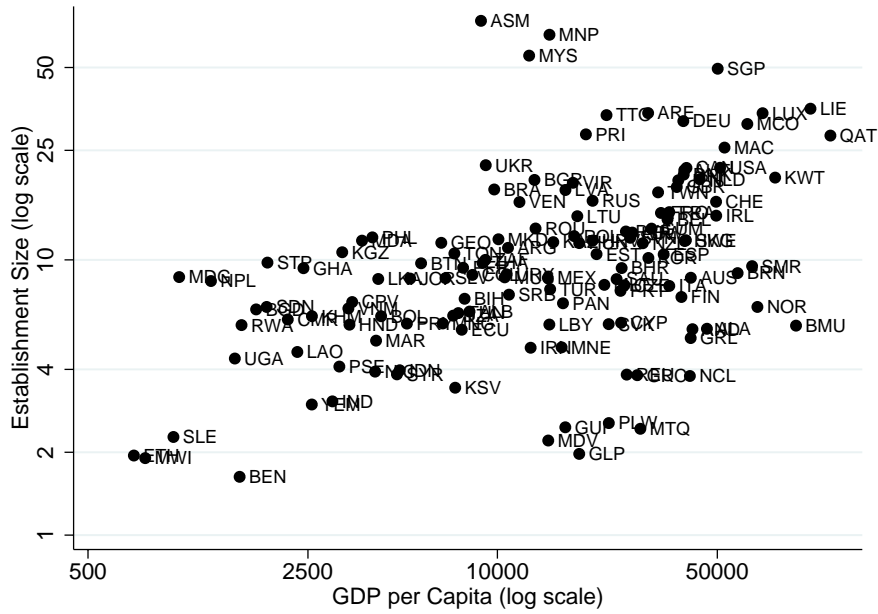
Figure 2 shows the average establishment size in non-agriculture (manufacturing and services) against GDP per capita for 91 countries. We calculate average size across manufacturing and services by combining our sectoral establishment-size measures with service and manufacturing shares from Duarte and Restuccia (2019), who use International Comparisons Program (ICP) data for 2005. If we denote sectoral shares in manufacturing and services by L_m and L_s , then average establishment size in non-agriculture is equal to $(L_m/size_m + L_s/size_s)^{-1}$. While 117 countries in our dataset have establishment size data for both manufacturing and services, only 91 of these have measures of sectoral shares in Duarte and Restuccia (2019). The elasticity of size (across sectors) with respect to GDP per capita is 0.33. This relationship is almost identical to that in Figure 1 because the corresponding elasticities are similar across sectors.⁴

Table 4 reports the elasticity of establishment size with respect to GDP per capita for each individual service industry, as well as for the more aggregated service and manufacturing sectors. In every industry, establishment size increases with development. Although the income elasticity of average size varies across disaggregated service industries, it does not vary systematically with relative average size differences across industries. Even though the income elasticity of average establishment size across all services is similar to the manufacturing elasticity of 0.3, the elasticity is as low as 0.13 for Wholesale services and as high as 0.45 for Information and Communication services.

Bartelsman et al. (2009) analyze firm size distributions across employers (firms with paid employees) in a smaller sample of countries, and suggest that much of the cross-country differences in average employment size are due to differences within sectors, rather than to differences in sectoral composition. Although our data includes a larger set of countries and the universe

⁴We also note that the share of services in non-agriculture is not systematically related to GDP per capita in our sample of countries. This fact is different from the systematic relationship between the share of agriculture or services and income per capita in the process of structural transformation, see for instance Duarte and Restuccia (2010) and Herrendorf et al. (2014).

Panel A: Manufacturing



Panel B: Services

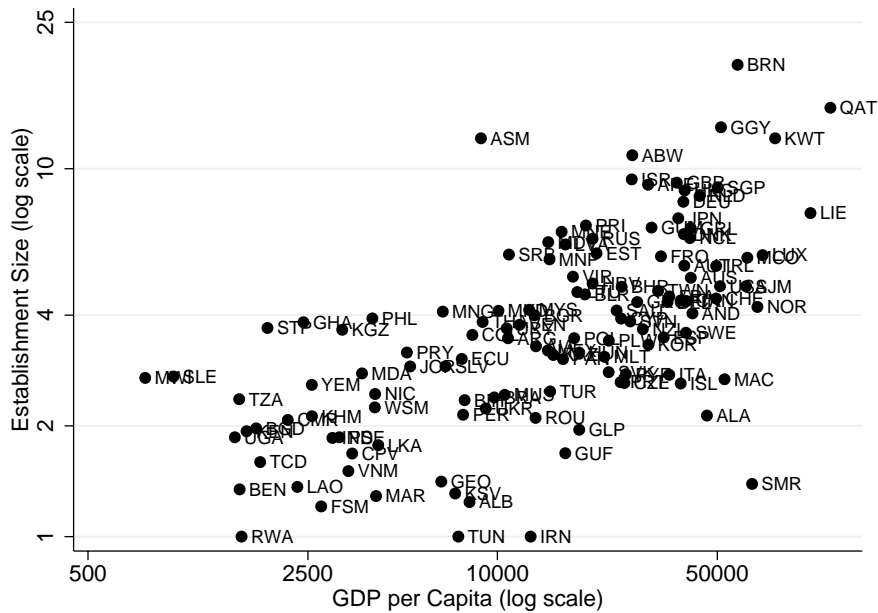


Figure 1: Establishment Size and GDP per Capita

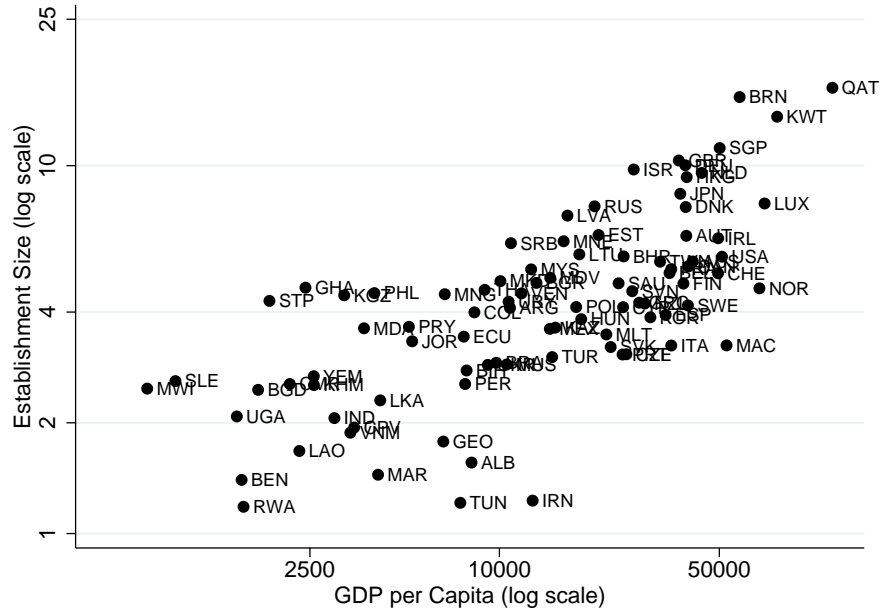


Figure 2: Establishment Size across Sectors and GDP per Capita

Table 4: INCOME ELASTICITY OF ESTABLISHMENT SIZE BY INDUSTRY

Services	0.32
Wholesale	0.13
Retail	0.39
Automobile Related	0.20
Accommodations and Restaurants	0.32
Transportation and Storage	0.28
Information and Communication	0.45
Real Estate and Business Services	0.15
Art, Culture, and Recreation	0.37
Finance and Insurance	0.22
Manufacturing	0.30

of firms (employers and nonemployers) in each country-sector, our results are consistent with theirs. Further, although countries obviously differ in sectoral composition, our results suggest that sectoral composition does not differ systematically in such a way as to lower average size in poor countries. Rather, in poorer countries firms are smaller in every sector.

2.3 Aggregation and Imputation

Our measures of average establishment size in service industries for many countries are calculated using the imputation procedure detailed previously. We evaluate the importance of these imputations for our findings by reporting elasticities using just the raw data. To do this, we use our preferred measure of size—persons engaged per establishment—for country-industries that report both persons engaged and the number of establishments. For country-industries without this data, we use persons engaged per firm. For countries without measures of persons engaged, we use employees per establishment. And for countries without measures of persons engaged or establishments, we use employees per firm. For each regression, we include dummy variables for the use of employees and for the use of firms in our size measure, as well as an additional dummy indicating whether ‘nonemployers’ (businesses without paid employees) are counted. We then regress (logged) size on (logged) GDP per capita separately for each industry. The estimated elasticities are reported in Table 5 along with our baseline elasticities from Table 4. The elasticities generated from the raw data are fairly close to those in our baseline data, within a few percentage points and preserving the direction and relative ranking of differences.

In calculating average establishment size for the service sector as a whole in each country, we also perform an imputation for countries that do not report data for some service industries. This could potentially bias the implied relationship between average size and development if for instance poorer countries tend to publish data only for industries with smaller establishments. Further, a size elasticity of income for the entire service sector may reflect both the size elasticity of income for each industry and systematic differences in industry composition across different levels of development. To address these concerns, we use the raw data and regress average

Table 5: INCOME ELASTICITY OF ESTABLISHMENT SIZE BY SERVICE INDUSTRY, BASELINE AND RAW DATA

	Baseline	Raw
Wholesale	0.13	0.13
Retail	0.39	0.42
Automobile Related	0.20	0.26
Accommodations and Restaurants	0.32	0.36
Transportation and Storage	0.28	0.30
Information and Communication	0.45	0.52
Real Estate and Business Services	0.15	0.16
Art, Culture, and Recreation	0.37	0.40
Finance and Insurance	0.22	0.24

size at the country-industry level on GDP per capita, again including dummies for the use of employees, firms, and nonemployers in the reported data, but now we use a pooled regression with dummy variables for each industry. Our estimated elasticity coefficient for GDP per capita is 0.32, identical to the elasticity we obtain using our baseline measure of average size for the entire service sector. We also perform the pooled data regression controlling for fixed industry effects separately for countries that report persons engaged per establishment, persons engaged per firm, employees per establishment, and employees per firm. We obtain the following corresponding coefficients, all significant at the 1% level: 0.37, 0.22, 0.41, and 0.38.

To summarize, the raw data confirms our main findings of a strong positive relationship between average establishment size and development.

3 Determinants of Establishment Size

Our comprehensive dataset of average establishment sizes across sectors and countries provides a unique opportunity to evaluate models of development along the establishment-size dimension. In this section, we take a first step in this direction by documenting the empirical relationship between average establishment size in each sector and relevant measures of country-specific

variables such as GDP per capita, sectoral employment, openness to trade, the share of external finance to GDP, firing costs, and firm-level distortions. We then discuss these relationships in the context of specific theories of development, providing motivation for the model we consider in the next section.

3.1 Data

We collect relevant cross-country data from several sources. Real GDP per capita is in purchasing power parity terms and is mainly from the Penn World Tables v8.0 (PWT) but also other sources as discussed in Section 2.1. Openness to trade is from the PWTv8.0, calculated as the value of imports plus exports as a share of GDP. Sectoral employment is the number of persons engaged in a sector. Specifically, we use the sectoral shares of manufacturing and services in [Duarte and Restuccia \(2019\)](#) multiplied by population from the PWTv8.0. We use population rather than total persons engaged due to a lack of total engaged data for many countries. Note that the sectoral shares in [Duarte and Restuccia \(2019\)](#) refer to the share of expenditures in manufacturing or services relative to total expenditures in manufacturing and services expressed in domestic prices. This share maps into employment shares in standard models of structural transformation. We use this data since it provides the largest country coverage. Nevertheless, the results are nearly identical using actual employment data across sectors from the 10-sector database for a restricted set of countries with available data. External finance measures the aggregate level of firm-level investment not financed internally, relative to GDP. We use a measure of external finance from [Buera et al. \(2011\)](#), calculated using data from [Beck et al. \(2010\)](#). Firing costs are from the OECD's Indicators of Employment Protection Legislation for 2008, which account for both individual and collective employee dismissals, and from [Heckman and Pagés \(2004\)](#). Firing costs measures the cost to a firm of firing a worker, both monetary and non-monetary (such as mandatory minimum notice before dismissal). To combine the different measures of firing costs, we regress the (logged) measures from the OECD on those from [Heckman and Pagés \(2004\)](#), then construct predicted measures consistent with

the OECD data for countries included in [Heckman and Pagés \(2004\)](#) but not included in the OECD data.

For firm-level distortions, rather than constructing broad measures of distortions based on dispersion in marginal products across firms, we focus on a narrower measure that emphasizes the systematic component of distortions, the extent to which distortions relate to establishment-level productivity or correlated distortions. We measure the systematic component of distortions by the elasticity of distortions with respect to productivity across firms. This is a more robust measure of misallocation and more closely related to firm decisions such as entry and investment which affect establishment size ([Hsieh and Klenow, 2014](#); [Bento and Restuccia, 2017](#)). We construct sectoral measures of correlated distortions for each country using establishment-level data from the World Bank’s Enterprise Surveys (WBES). The WBES dataset includes data from mostly low- and middle-income countries collected through face-to-face surveys, and contains standardized information about sales, intermediate purchases, inputs, and a host of other variables for establishments in over 100 countries for at least one year from 2002 to 2012. We back out our measure of establishment-level distortions and productivity for each establishment within a country-industry-year by imposing the following structure on the data, which is consistent with the model we develop in Section 4. We assume establishments within an industry differ with respect to their productivity Z , and produce output according to the following decreasing returns to scale technology;

$$y = Z\ell^\alpha, \tag{1}$$

where $\alpha \in (0, 1)$ and ℓ denotes labor. Note that we abstract from capital in our measures because this allows us to substantially increase the sample of establishments for each country and the number of countries in our analysis, as a large number of establishments in WBES do not report capital and sample sizes in WBES are small. Nevertheless, it is well-known that empirical measures of wedges for different factor inputs are highly correlated so they are well captured by a composite output wedge. Assuming establishments take the price of output p and

the wage w as given, each establishment optimally chooses ℓ such that its after-tax marginal revenue product is equal to the prevailing wage w . This suggests the following relationship between an establishment’s labor productivity and its distortion (which we model as an explicit tax on output τ);

$$\frac{py}{\ell} = \frac{1}{(1-\tau)} \left(\frac{w}{\alpha}\right) \propto \frac{1}{(1-\tau)},$$

where py is an establishment’s value added (sales minus intermediate inputs). Following [Hsieh and Klenow \(2009\)](#), we use an establishment’s total wage bill (including benefits) in our computations instead of employment in order to control for differences in human capital across establishments. Our measure of labor productivity for each establishment is calculated *relative to* the weighted average of labor productivity across all establishments within the same industry, weighted by each establishment’s share of value added. Establishment-level productivity (denoted by Z) is inferred under the same assumptions using the model-implied relationship between an establishment’s revenue and its output:

$$Z = \frac{y}{\ell^\alpha} \propto \frac{py}{\ell^\alpha}.$$

Following [Hsieh and Klenow \(2009\)](#), productivity is calculated relative to industry productivity in the absence of distortions, $[\text{mean}(Z^{1/(1-\alpha)})]^{1-\alpha}$. Note that the above calculations require a value for α . We use a value of 2/3 to be comparable to [Hsieh and Klenow \(2014\)](#) and [Bento and Restuccia \(2017\)](#).

We estimate the elasticity of distortions with respect to productivity across establishments in each sector using the above measures by running an OLS regression of logged distortions on logged productivity across establishments. Following [Hsieh and Klenow \(2014\)](#) we trim the 1 percent tails of both distortions and productivity for each country-industry to remove outliers. Some countries have data for two or more years, so we average elasticities over all years, weighting by the number of observations in each year. We obtain elasticities for 74 countries in the manufacturing sector and for 63 countries in the service sector. In manufacturing,

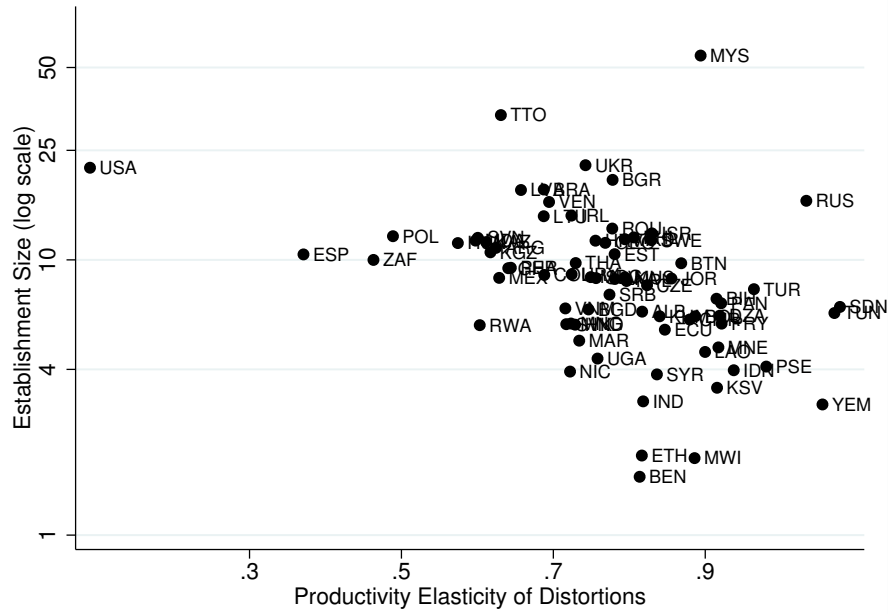
elasticities range from 0.37 to 1.08, averaging 0.78. In services, elasticities range from 0.36 to 1.09, averaging a higher 0.87. Countries with a higher productivity elasticity of distortions in manufacturing also tend to have a higher elasticity in services. The corresponding correlation coefficient between the two sectoral measures of distortions is 0.62.

Figure 3 illustrates how correlated distortions vary with average establishment size in each sector. In both manufacturing and services, average size is lower when the productivity elasticity of distortions is higher. Although the United States is not included in the WBES data, we include it in Figure 3 for comparison, using an elasticity of 0.09 for both manufacturing and services. Hsieh and Klenow (2014) estimate a productivity elasticity of distortions of 0.09 for U.S. manufacturing and we assume the same value for the U.S. service sector. Our measures of distortions for each country-sector depend in part on the value of α . For robustness, we also re-estimate the elasticities using a higher value of $\alpha = 0.8$ (rather than $2/3$). The resulting relationships between the productivity elasticity of distortions and average size are similarly negative and significant, both in manufacturing and in services. Our assumption that α is constant across sectors is consistent with Herrendorf and Valentinyi (2008), who find relatively constant factor shares across sectors in the U.S. data. Our assumption that α is constant across countries is consistent with Gollin (2002), who finds that the labor share of income is not systematically related to GDP per capita across countries.

3.2 Theories of Development

In order to provide evidence on the potential determinants of average establishment size across sectors and countries, we estimate reduced-form empirical relationships between establishment size and several indicator variables at the core of prominent theories of productivity and development. Tables 6 and 7 document the estimated coefficients from bivariate and multivariate regressions of average establishment size on subsets of country-specific variables such as GDP per capita, openness to trade, firing costs, external finance, and firm-level distortions. We restrict multivariate regressions to subsets of explanatory variables for which we have at least 30

Panel A: Manufacturing



Panel B: Services

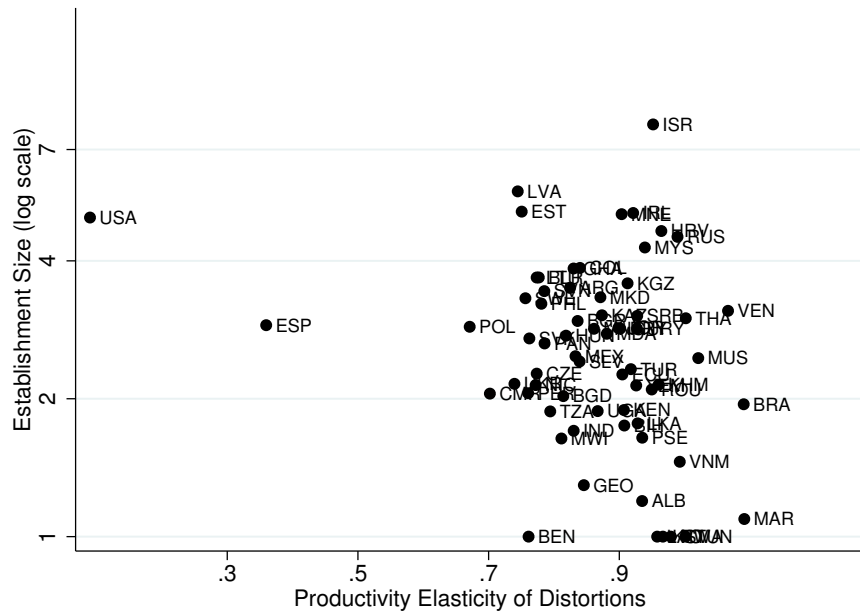


Figure 3: Establishment Size and Productivity Elasticity of Distortions

observations. Table 8 reports the results of similar regressions for the average establishment size ratio between manufacturing and services. We discuss these empirical results in the context of prominent models of development and aggregate productivity.

Entry Costs A number of papers, including [Barseghyan and DiCecio \(2011\)](#), [Moscoso Boedo and Mukoyama \(2012\)](#), and [Bento \(2014b\)](#), consider entry costs (the cost of starting a firm) as a potential source of low productivity in poor countries. Entry costs reduce entry and the number of firms in equilibrium, increasing the size of firms above their optimal size and thereby lowering aggregate productivity. There is no good data on entry costs, for instance the often-used variable “startup cost” as a measure of regulatory entry costs from the World Bank’s Doing Business data indicates the cost of incorporating a large firm, not the cost of starting a firm. Nevertheless, the positive relationships between size and GDP per capita in Panel A of Tables 6 and 7 suggest that entry costs are not in fact higher in poor countries.

A related issue is the modeling of entry costs (and other fixed costs) in quantitative models. Fixed costs are often specified either in terms of goods or in terms of a fixed factor such as labor. The units in which costs are specified has implications for how various policies affect aggregate outcomes. It also has implications on how growth affects average establishment size over time. If wages grow along with GDP per capita, then a cost of entry in units of labor grows proportionately, which typically implies a constant number of firms per capita and average firm size. Instead, a goods cost of entry in a growing economy shrinks relative to the operating profits of firms, implying more entry and lower average firm size. [Bollard et al. \(2016\)](#) examine firm-level panel data from the United States, China, and India, and conclude that average firm size does not shrink over time in growing economies, suggesting that fixed costs should be specified in terms of a fixed factor such as labor rather than (purely) goods. The positive coefficients on GDP per capita in Tables 6 and 7 suggest that specifying fixed costs in terms of goods generates counter-factual implications with respect to the cross-country data as well. Establishments are definitely not smaller in rich countries in manufacturing or services.

Table 6: DETERMINANTS OF AVERAGE ESTABLISHMENT SIZE IN MANUFACTURING

Panel A: Bivariate regressions						
Dependent variable: Average establishment size in manufacturing						
Independent variables:						
GDP per capita	0.30***					
	(0.05)					
Mnfg. employment		0.03				
		(0.05)				
External financing			0.42***			
			(0.06)			
Firing costs				-0.31***		
				(0.08)		
Openness to trade					0.41***	
					(0.11)	
Correlated distortions (Mnfg.)						-1.52***
						(0.33)
Countries	134	74	56	52	95	75
R^2	0.22	0.01	0.43	0.15	0.19	0.15

Panel B: Multivariate regressions						
Dependent variable: Average establishment size in manufacturing						
Independent variables:						
External financing	0.21*	0.38***	0.29**			
	(0.11)	(0.07)	(0.11)			
Firing costs	-0.17*			-0.29***	-0.20*	
	(0.10)			(0.08)	(0.12)	
Openness to trade		0.17		0.16		0.30**
		(0.13)		(0.12)		(0.13)
Correlated distortions (Mnfg.)			-1.33***		-0.85**	-1.34***
			(0.42)		(0.39)	(0.39)
Countries	38	50	34	47	32	60
R^2	0.26	0.50	0.43	0.20	0.26	0.25

Notes: All variables logged, except for correlated distortions. See the text for the definition of variables and sources. Robust standard errors in parentheses. ***, **, and * refer to one, five, and ten percent levels of significance.

Table 7: DETERMINANTS OF AVERAGE ESTABLISHMENT SIZE IN SERVICES

Panel A: Bivariate regressions						
Dependent variable: Average establishment size in services						
Independent variables:						
GDP per capita	0.32***					
	(0.04)					
Services employment		-0.01				
		(0.03)				
External financing			0.33***			
			(0.06)			
Firing costs				-0.32***		
				(0.07)		
Openness to trade					0.38***	
					(0.09)	
Correlated distortions (Serv.)						-0.84***
						(0.29)
Countries	127	82	53	49	100	66
R^2	0.39	0.00	0.40	0.27	0.22	0.07

Panel B: Multivariate regressions				
Dependent variable: Average establishment size in services				
Independent variables:				
External financing	0.20***	0.25***		
	(0.06)	(0.08)		
Firing costs	-0.20***		-0.31***	
	(0.07)		(0.07)	
Openness to trade		0.20*	0.10	0.26***
		(0.11)	(0.09)	(0.10)
Correlated distortions (Mnfg.)				-0.74**
				(0.31)
Countries	37	53	48	63
R^2	0.43	0.44	0.30	0.21

Notes: All variables logged, except for correlated distortions. See the text for the definition of variables and sources. Robust standard errors in parentheses. ***, **, and * refer to one, five, and ten percent levels of significance.

Table 8: DETERMINANTS OF SIZE RATIO MANUFACTURING TO SERVICES

Dependent variable: Ratio of average size manufacturing to services						
Independent variables:						
GDP per capita	0.01					
	(0.05)					
Employment ratio (M/S)		0.30				
		(0.33)				
External financing			0.11			
			(0.08)			
Firing costs				0.03		
				(0.09)		
Openness to trade					0.04	
					(0.09)	
Correlated distortions gap (M-S)						-1.16***
						(0.40)
Countries	117	91	50	48	95	63
R^2	0.00	0.01	0.05	0.00	0.00	0.10

Notes: All variables logged, except for correlated distortions. See the text for the definition of variables and sources. Robust standard errors in parentheses. *** refers to a one percent level of significance.

Financial Constraints Buera et al. (2011, 2015) develop what has since become a workhorse model of heterogeneous firms with financial constraints. In the one-sector version of the model, the effect of financial constraints on average firm size is theoretically ambiguous. Financial constraints lead to slower firm growth, which encourages entry due to lower average productivity among incumbents (Atkeson and Burstein, 2010). But potential entrepreneurs must accumulate savings in order to self-finance fixed costs, which lowers entry. In the two-sector version of the model, where sectors are differentiated with respect to the size of fixed costs (larger in manufacturing than in services), Buera et al. (2011) specification of fixed costs in terms of sectoral output results in less ambiguous predictions. Financial constraints bite harder in manufacturing, as potential entrepreneurs must save longer to start a (relatively) large manufacturing firm. As a result, the relative price of manufacturing increases. This relative change creates an additional wedge between the ease of entry in manufacturing and services, as the fixed cost of manufacturing output becomes more costly. The result is a larger average size in manufacturing, and a smaller average size in services.

Tables 6 and 7 document a positive relationship between average size and the extent of external financing in an economy, both for manufacturing and service establishments. The extent of external financing is a widely-used proxy for financial constraints, as the theory predicts a monotonically negative relationship between the extent of constraints and external financing. The results in Table 6 are inconsistent with the prediction of the model that manufacturing firms should be larger with tighter financial constraints and less external financing. Moreover, Table 8 suggests the ratio of average size in manufacturing to that in services is uncorrelated with financial constraints. Table 4 also suggests that average size differences across industries are uncorrelated with GDP per capita. To the extent that financial constraints are more of a problem in poor countries, the disaggregated industry data also seems inconsistent with the theory. It is possible that a specification of fixed costs in terms of labor would help reconcile the implications of the model with data.

Firing Costs Hopenhayn and Rogerson (1993) extend the model of firm dynamics in Hopenhayn (1992) to evaluate the general equilibrium impact of firing costs on aggregate employment. Hopenhayn and Rogerson (1993) find that firing costs dampen the response of firms' labor decisions to productivity shocks, lowering aggregate employment and reducing aggregate TFP due to a misallocation of labor across firms and fewer firms in equilibrium. Average firm size is predicted to increase with firing costs. This prediction is inconsistent with the data as suggested by the results in Tables 6 and 7. Among economies with data on firing costs, higher firing costs are associated with smaller establishments, both in services and manufacturing. Although the coefficients are not significant across all specifications, they are all negative. Note that Hopenhayn and Rogerson (1993) predict a quantitatively small effect on firm size, which we might not pick up with so few observations and controls. But the negative relationship between firing costs and establishment size suggests that more research into the mechanisms through which firing costs operate is needed. One mechanism that we highlight below is that the misallocation generated by firing costs may induce a change in establishment-level productivity, potentially

generating implications for average establishment size more in line with data.⁵

Endogenous Markups Several strands of the development literature have incorporated endogenous markups in models of firm size. [Melitz and Ottaviano \(2008\)](#) and [Desmet and Parente \(2010\)](#), for example, use endogenous markups to investigate the pro-competitive gains from trade, while [Bento \(2014a\)](#) incorporates endogenous markups in a model of innovation in order to rationalize the empirical relationship between competition and innovation. To our knowledge, almost all models with endogenous markups and endogenous firm size imply a positive relationship between population and firm size.⁶ In models with constant markups such as [Melitz \(2003\)](#), an increase in population results in a proportional increase in the number of firms, leaving average firm size unchanged. But if markups are endogenous and depend on the number of competitors, more firms leads to lower markups, thereby lowering profits (relative to the constant markup case). As a result, the number of firms increases less than population, and average firm size is larger. Panel A of Tables 6 and 7 show that establishment size in services and manufacturing are unrelated to sectoral employment (the product of population and the sectoral share of employment). Although not shown, average establishment size in each sector is also unrelated to population. Interpreted through the lens of models with endogenous markups, these results suggest that while larger populations are associated with more firms, they are not associated with more competition.

Trade Models of international trade imply that the relationship between openness to trade and firm size depends on which trade barrier is driving the variation in openness—tariffs vs. fixed trade costs ([Melitz, 2003](#)), fixed firm-specific vs. product-specific trade costs ([Bernard et al., 2011](#)). For example, lowering tariffs across countries should cause exporters to expand while firms who only sell domestically contract or exit in response to the increase in imports,

⁵See [Da-Rocha et al. \(2016\)](#) and [Da-Rocha et al. \(2017\)](#) for frameworks where firing costs and misallocation induce changes in the productivity distribution.

⁶An exception is [Bento \(2018b\)](#) where firms face a market-entry cost which is increasing in the number of markets entered. If the number of markets in an economy increases with population, then so does the number of firms, but each firm continues to serve the same number of markets. As a result, the number of firms in each market is independent of population.

resulting in larger firms. Table 6 reports a positive relationship between establishment size in manufacturing and openness to trade, consistent with variation in tariffs (and distance-related costs) as the dominant driving force behind differences in the extent of trade across countries. But this result is not robust across all specifications in Table 6. Further, Table 7 reports an identical unconditional relationship between openness to trade and average establishment size in services, and Table 8 shows that the relative size of establishments in manufacturing relative to services is unrelated to how open a country is to trade. To the extent that goods are more likely to be traded than services, the data seems inconsistent with tariffs driving average size. Rather, the data seems consistent with differences in openness to trade across countries being driven by variation in several types of trade barriers, with no one barrier dominating empirically. It is possible that incorporating inter-sectoral linkages and trade in services, which has been emphasized less in the trade literature, may generate richer implications for relative establishment sizes from trade.

Correlated Distortions A recent but growing literature evaluates the impact of misallocation when the wedges faced by firms are dependent on firm characteristics that are in part chosen by firms. For instance, [Hsieh and Klenow \(2014\)](#), [Bento and Restuccia \(2017\)](#), and [Guner et al. \(2018\)](#) consider models of misallocation where firms face distortionary effective taxes that depend on firm productivity, and firms take this into account when choosing their productivity. In addition to the impact on aggregate productivity from the misallocation of inputs across firms, correlated distortions reduce the marginal incentive of firms to invest in productivity, further reducing aggregate productivity and lowering average firm size. The empirical results in Tables 6, 7, and 8 are consistent with this broad theory. Average establishment size in each sector is declining in the extent of correlated distortions, and the ratio of average size in manufacturing relative to services is decreasing in the difference between correlated distortions in the two sectors. Note that the type of misallocation (correlated distortions) considered in this literature comprises many types of policies and institutions that are known to create the specific patterns including informality, financial frictions, selective regulation, firing costs, and

trade policy, among others. For example, financial frictions may disproportionately impact the more productive establishments as the credit friction is likely to more greatly affect their optimal size, thus reducing the return to investment in productivity and decreasing the average size of establishments. In calculating the extent of correlated distortions using the reported decisions of establishments, our measure of correlated distortions captures the net impact of many policies and institutions working through this mechanism.

We note that correlated distortions arise from specific policies and institutions discussed earlier such as financial frictions, firing costs, and trade restrictions (Buera et al., 2013; Hopenhayn, 2014). As a result, an important element in reconciling the implications of models featuring correlated distortions with the empirical evidence is the endogenous determination of firm-level productivity. The mechanism generating a link between correlated distortions and average size relies on the property that establishments reduce their investment in productivity to avoid higher effective taxes, that is, correlated distortions effectively reduce the return to productivity investment. This lowers the share of revenue allocated to investment, which results in higher profits, encouraging entry and thereby lowering average establishment size in equilibrium (Hopenhayn, 2016).

We caution that our empirical results in Tables 6, 7, and 8 cannot be used to assess which variables are most important in explaining the variance in average size across countries since country samples vary in each regression. For instance, when restricting to the same sample of countries, financial frictions and correlated distortions are equally important in accounting for the variance in average establishment sizes across countries, which combined account for about 50 percent of the variance.

To summarize, models of correlated distortions (including financial frictions, firing costs, and trade restrictions) with endogenous entry and productivity generate qualitative predictions consistent with the establishment size data across sectors and countries. Appendix B shows that the empirical determinants of establishment size just discussed are robust to instead using only the raw data.. In the next section, we incorporate correlated distortions into a two-sector

model to assess the impact on sectoral average establishment size, and to quantify the impact of correlated distortions on sectoral and aggregate productivity.

4 A Model of Size across Sectors and Countries

We develop a two-sector model of establishment size with endogenous sectoral employment shares. We focus on the non-agricultural sector and hence the environment includes manufacturing and services. We abstract from agriculture as the details of the production unit and the relevant definition of size are different for agriculture than for the rest of the economy. Nevertheless, the same economics emphasized for non-agricultural establishments applies for agriculture as well ([Adamopoulos and Restuccia, 2014](#)).

There is a stand-in household (of measure one) endowed with one unit of productive time, and there are two goods produced in sectors: manufacturing and services. The production unit in each sector is an establishment operating with a decreasing returns to scale technology.⁷ There are a large number of potential firms who are free to enter, but must pay a fixed entry cost and make a costly productivity-investment decision before producing. Producers face output distortions which may be correlated with firm-level productivity, and entrants take these policy distortions into account when investing in productivity. We assume a static economy for simplicity, but our results do not depend critically on this assumption. We study a competitive equilibrium of the economy in which firms take prices as given and free entry ensures the value of entry is driven to zero. We then consider how the extent of correlated distortions affects the number of firms, investment, and aggregate output. We begin by describing the environment in more detail.

⁷Throughout we use the terms ‘firm’, ‘producer’, and ‘establishment’ interchangeably.

4.1 Economic Environment

Preferences and Endowments There is a representative household of measure 1 that has preferences for manufacturing (m) and service (s) goods according to the following utility function,

$$U(C_m, C_s) = [bC_m^\rho + (1 - b)C_s^\rho]^{1/\rho}, \quad b \in (0, 1), \quad \rho < 0. \quad (2)$$

The household is endowed with one unit of productive time each period that is supplied inelastically to the market.

Technologies In each sector, an homogeneous good is produced in establishments. An establishment produces output y according to the following production function,

$$y(h, z, \ell; i) = A_i h z \ell^\alpha,$$

where ℓ is labor, A_i is a sector-wide productivity term, h and z are idiosyncratic productivity terms, and $\alpha \in (0, 1)$. Of the productivity terms, h is chosen by the firm upon entering the market with a sector-specific labor cost $w c_h^i h^\theta$ with $c_h^i > 0$ and where w denotes the real wage; and z is exogenous and known after entering and investing in h according to a distribution with *cdf* $F(z)$. There are a large number of potential firms, each of which can become a producer by incurring a sector-specific labor cost of entry $w c_e^i$ with $c_e^i > 0$.

Market Structure and Distortions Consider a competitive economy where households and firms take prices as given. There are a large number of firms operating in the manufacturing and services sectors, renting labor services from consumers at the rate w and selling consumption goods to households at prices p_m and p_s . In each sector, producers face a proportional tax on revenue τ which depends on firm-level productivity in the following way,

$$1 - \tau(h, z; i) = \kappa_i \cdot (hz)^{-\gamma_i}, \quad \gamma_i \geq 0. \quad (3)$$

where γ_i is the elasticity of distortions with respect to productivity in sector i and κ_i is a level shifter that changes with γ_i to ensure that the sectoral average weighted distortion (weighted by each firm's sectoral revenue share) is constant across different levels of γ_i . This allows us to isolate the impact of a higher productivity elasticity of distortions, keeping average distortions constant. Considering different levels of γ_i while keeping the average distortion constant is consistent with the counterfactuals examined in [Hsieh and Klenow \(2009\)](#) and [Bento and Restuccia \(2017\)](#). In what follows, we assume that this average tax is equal to zero.

Definition of Equilibrium A *competitive equilibrium* is a list of prices (w, p_m, p_s) ; decision functions for firms in each sector: labor demand $\ell(h, z)$, operating profits $\pi(h, z)$, endogenous productivity h_i , value of entry V_e^i , and number of firms N_i in each sector; and allocation for consumers: C_m, C_s such that;

- (i) Given w, p_s , and p_m ; allocation (C_m, C_s) solves the household problem, i.e. maximizes utility subject to the budget constraint.
- (ii) Given w, p_s , and p_m ; producers in each sector choose labor $\ell(h, z)$ to maximize per-period operating profits $\pi(h, z)$, which in turn determine the value of an incumbent producer.
- (iii) Given w, p_s , and p_s ; entrants choose productivity h_i in each sector to maximize the expected value of operating profits net of investment.
- (iv) Free entry $V_e^i = 0$ in each sector ensures that the expected value of operating profits is equal to the optimal productivity investment plus the entry cost.
- (v) Markets clear, i.e., the supply of labor (equal to one) is equal to the quantity of labor demanded by firms plus the entry and investment costs of labor; and the supply and demand of the output good are equal.

4.2 Characterization

The household problem is simply to maximize utility subject to the budget constraint by choosing consumption in each sector,

$$\max [bC_m^\rho + (1-b)C_s^\rho]^{1/\rho},$$

subject to

$$p_m C_m + p_s C_s = w.$$

The solution to this problem (C_m, C_s) satisfies,

$$\frac{b}{1-b} \left(\frac{C_m}{C_s} \right)^{\rho-1} = \frac{p_m}{p_s}, \quad (4)$$

and the budget constraint specified above.

A producer with productivity hz in sector i solves the following problem,

$$\max_{\ell} (1 - \tau(h, z; i)) p_i y(h, z, \ell; i) - w\ell,$$

where $y(h, z, \ell; i) = A_i h z \ell^\alpha$ and $(1 - \tau(h, z; i)) = \kappa_i \cdot (hz)^{-\gamma_i}$, which implies the following optimal output $y(h, z; i)$ and operating profits $\pi(h, z; i)$,

$$y(h, z; i) = (A_i h z)^{\frac{1}{1-\alpha}} \left(\frac{\alpha p_i \cdot (1 - \tau(h, z; i))}{w} \right)^{\frac{\alpha}{1-\alpha}}, \quad (5)$$

$$\pi(h, z; i) = \left(\frac{\alpha^\alpha p_i \cdot (1 - \tau(h, z; i)) A_i h z}{w^\alpha} \right)^{\frac{1}{1-\alpha}} (1 - \alpha). \quad (6)$$

Before learning z and producing, an entrant to sector i must choose h to solve the following problem,

$$\max_h \mathbb{E}[\pi(h, z; i)] - w c_h^i h^\theta,$$

which implies

$$wc_h^i h_i^\theta = \left(\frac{1 - \gamma_i}{(1 - \alpha)\theta} \right) \mathbb{E}_z[\pi(z; i)]. \quad (7)$$

Finally, free entry ensures that firms enter each sector until expected operating profits net of productivity investment is exactly equal to the cost of entry,

$$wc_e^i = \left(\frac{(1 - \alpha)\theta - 1 + \gamma_i}{(1 - \alpha)\theta} \right) \mathbb{E}_z[\pi(z; i)]. \quad (8)$$

Note that equations (7) and (8) together imply the following h is chosen by firms in each sector,

$$h_i = \left(\frac{c_e^i(1 - \gamma_i)}{c_h^i[(1 - \alpha)\theta - 1 + \gamma_i]} \right)^{\frac{1}{\theta}}. \quad (9)$$

The average (weighted) distortion in each sector is characterized by the following expression,

$$1 - \bar{\tau}_i = \frac{\mathbb{E}_z[y(z; i)(1 - \tau(z; i))]}{\mathbb{E}_z[y(z; i)]}. \quad (10)$$

To keep $\bar{\tau}_i$ equal to zero across different values of γ_i , we set κ_i equal to

$$\kappa_i = h_i^{\gamma_i} \frac{\mathbb{E}_z \left[z^{\frac{1 - \alpha \gamma_i}{1 - \alpha}} \right]}{\mathbb{E}_z \left[z^{\frac{1 - \gamma_i}{1 - \alpha}} \right]}. \quad (11)$$

Combining equations (5), (6), (8), and (11), we can now derive the average size of firms in each sector. If N_i and L_i denote the number of firms and the quantity of labor in sector i , then average firm size in sector i is

$$\frac{L_i}{N_i} = \frac{\theta c_e^i}{(1 - \alpha)\theta - 1 + \gamma_i}. \quad (12)$$

We note that average size in each sector is increasing in entry costs c_e^i and decreasing in γ_i . Average size is *not* dependent on exogenous sectoral productivity. Average size is also independent of exogenous aggregate productivity.

Equations (5), (12), labor-market clearing, and goods-market clearing can be combined to

derive measured TFP in each sector;

$$TFP_i \equiv \frac{C_i}{L_i} = \frac{\alpha^\alpha}{\theta^{1-\alpha}} A_i \left(\frac{1-\gamma_i}{c_h^i} \right)^{\frac{1}{\theta}} \left[\frac{(1-\alpha)\theta - 1 + \gamma_i}{c_e^i} \right]^{\frac{(1-\alpha)\theta-1}{\theta}} \frac{\mathbb{E}_z \left[z^{\frac{1-\alpha\gamma_i}{1-\alpha}} \right]}{\mathbb{E}_z \left[z^{\frac{1-\gamma_i}{1-\alpha}} \right]^\alpha}. \quad (13)$$

Sectoral TFP is increasing in exogenous productivity A_i , and decreasing in the costs of entry and productivity investment (c_e^i and c_h^i) as well as the productivity elasticity of distortions γ_i . The first two bracketed terms in the TFP_i expression in equation (13) represent the combined effects of (c_e^i, c_h^i, γ_i) on the number of firms per worker N_i and the endogenous productivity term h_i , which affect TFP_i in opposite ways. The last term in equation (13) represents the negative effect on aggregate TFP_i arising from misallocation due to distortions γ_i .

The price of manufactured goods, relative to services, is

$$\frac{p_m}{p_s} = \frac{TFP_s}{TFP_m}. \quad (14)$$

We solve for sectoral labor shares by combining the above results with equation (4) from the household's problem. The shares of labor allocated to each sector are:

$$L_m = \frac{\Psi}{\Psi + 1}; \quad L_s = \frac{1}{\Psi + 1}, \quad (15)$$

where

$$\Psi \equiv \left(\frac{b}{1-b} \right)^{\frac{1}{1-\rho}} \left(\frac{TFP_m}{TFP_s} \right)^{\frac{\rho}{1-\rho}}.$$

In the Cobb-Douglas case, $L_m = b$ when ρ goes to 0, and sectoral shares do not depend on sectoral productivities. When $\rho < 0$, a higher ratio of productivity in manufacturing to services results in a higher share of labor in services. In a dynamic version of the model, if labor productivity growth is faster in manufacturing than in services, then labor reallocates over time from manufacturing to services.

We define real GDP as equivalent to $U(C_m, C_s)$. If the price of a unit of real GDP is normalized

to one, then GDP is equal to the wage w . Using equations (2), (13), (14), and (15), we can express the wage as

$$w = (\Psi + 1)^{-1} [b \cdot TFP_m^\rho \cdot \Psi^\rho + (1 - b) \cdot TFP_s^\rho]^{\frac{1}{\rho}},$$

or

$$w = (1 - b)^{\frac{1}{\rho}} \cdot TFP_s \cdot (1 + \Psi)^{\frac{1-\rho}{\rho}}. \quad (16)$$

5 Quantitative Analysis

We calibrate a benchmark economy to U.S. data. We use the calibrated model to quantify the implications of cross-country variations in the degree of correlated distortions and entry costs in each sector on establishment size and productivity in each sector; and on the sectoral composition of the economy and aggregate output.

5.1 Calibration

We calibrate a benchmark economy to U.S. data, with $\bar{\tau}_i$ set to zero and γ_i to 0.09 in each sector. The parameters to calibrate are: preferences (b, ρ) , entry and productivity investment costs (c_e^i, c_h^i) , technology (α, θ) , and exogenous productivity distributions $(A_i, F(z))$. We assume $c_h^i = 1$ in both sectors since our model cannot separately identify the level of costs c_h^i and exogenous sector productivity A_i .

We assume $\alpha = 2/3$ which maps into the curvature parameter used in the monopolistic competition model of [Hsieh and Klenow \(2009\)](#), where the elasticity of substitution between varieties is assumed to be 3. For θ , we choose this parameter to match the elasticity of revenue with respect to investment in productivity reported in [Bento and Restuccia \(2017\)](#), equal to

0.72. In our model, this elasticity is equal to

$$\frac{1 - \alpha\gamma_i}{(1 - \alpha)\theta}.$$

With the value of α and using $\gamma_i = 0.09$ in the U.S. in both sectors, we obtain $\theta = 3.92$. We choose c_e^i to match average size in manufacturing and services from equation (12) in the United States, obtaining $c_e^m = 2.18$ and $c_e^s = 0.50$. From [Duarte and Restuccia \(2010\)](#), we select $\rho = -1.5$. Given ρ and a target for the relative price of manufacturing to services, we obtain $b = 0.05$ to match U.S. sectoral shares from equation (15).

We obtain distributions of z for each sector in our benchmark economy by interpreting the employment-size distributions of establishments in the United States through the lens of our model. For this purpose we use data from the U.S. Census Bureau. Our model with no distortions implies a simple mapping between productivity z and employment in equation (5), such that the labor demand of establishment i relative to that of establishment j is;

$$\frac{\ell_i}{\ell_j} = \left(\frac{z_i}{z_j} \right)^{\frac{1-\gamma}{1-\alpha}}.$$

The U.S. data reports the number of establishments in each employment-size bin, using the number of paid employees as the measure of employment. We adjust the data to account for both unpaid workers and owner-managers by assuming the number of persons engaged among nonemployers ranges from 1-2, and by assuming the number of persons engaged among establishments with 1-4 paid employees ranges from 3-4. As a result, we obtain separate distributions of persons engaged per establishment in manufacturing and services ranging from 1 to 3000 persons. In each sector, we match exactly the fraction of establishments falling within each size bin, illustrated in Figure 4. The size distribution of establishments in services is more skewed to the right than in manufacturing, with less dispersion in size among service establishments. For our calibrated productivity distributions, we assume the productivity of establishments within each bin are uniformly distributed.

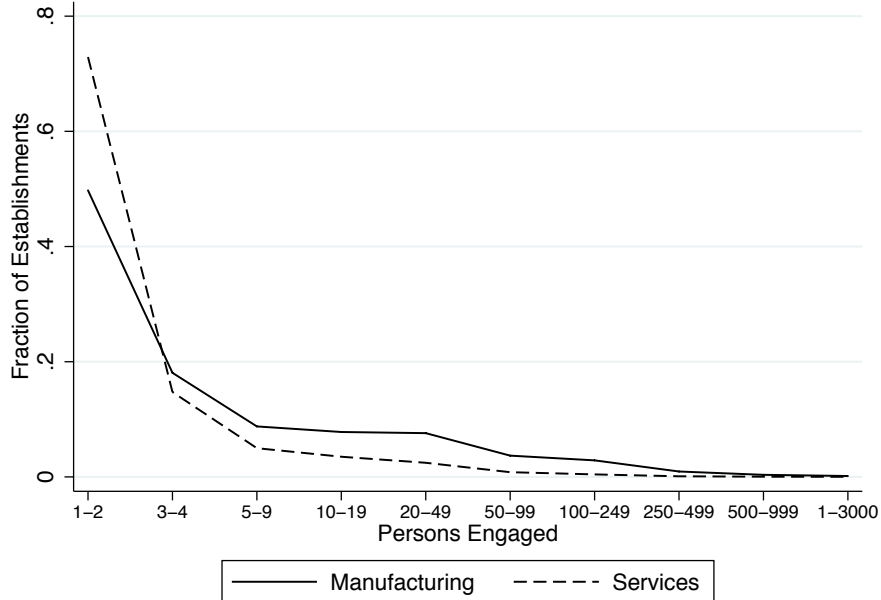


Figure 4: Establishment Size Distributions, U.S. Data

Importantly, we treat very small establishments (including ‘nonemployers’) as simply small establishments - that is, establishments with access to the same technology as larger establishments, but with lower productivity. This is consistent with a growing literature comparing the characteristics of nonemployers to employers in the U.S., including [Davis et al. \(2009\)](#), [Acs et al. \(2009\)](#), and [Fairlie et al. \(2018\)](#). The literature has documented that nonemployers: coexist with employers within distinct industries; share similar survival rates with small employer establishments; and transition into employer status at a rate consistent with average employment growth across small employers. Although we do not have cross-country data on the share of establishments that are nonemployers, in the next section we discuss the implications of our model for the share of nonemployers.

This completes the calibration for the benchmark economy. We now characterize parameters that differ across countries. We assume countries are otherwise identical to the benchmark economy except on 6 exogenous parameters: the productivity elasticity of distortions γ_i , entry costs c_e^i , and exogenous productivity A_i in each sector i . The values of γ_i for each sector and country are given by the data illustrated in [Figure 3](#). We choose c_e^i in each sector and country

to match the average establishment size data across sectors and countries reported in Figure 1. Combining the above values with sectoral shares in each country from Duarte and Restuccia (2019), we back out the ratio of exogenous productivity between manufacturing and services A_m/A_s using equations (13) and (15). Finally, differences in real GDP per capita determine the absolute differences in A_m (and A_s) across countries using equation (16).

5.2 Results

The Role of Correlated Distortions We start by reporting how variation in γ_i alone across countries affects average establishment size, sectoral productivity, and sectoral output. Table 9 reports the results of increases in γ , from 0.09 in the benchmark economy to 0.9 in the most distorted economy. For these results, we assume $\gamma_m = \gamma_s$ so that correlated distortions are the same across sectors within an economy but vary across economies. For instance, an increase in correlated distortions from 0.09 to 0.7—well within the range observed in the cross-country elasticities reported earlier—implies that average establishment size drops by about 60 percent in both sectors: from 22 persons engaged to 9 in the manufacturing sector and from 5 to 2 persons engaged in the service sector. This range in establishment size corresponds well with the range reported earlier in the data across rich and developing countries. Similarly, productivity in each sector drops by 41 percent, while the direct impact of misallocation is about half as large. As a result of these effects, sectoral output drops by about 35 percent in each sector with a similar effect on aggregate output.

Average Establishment Size We now study the implications of our data on the productivity elasticity of distortions on average establishment size across sectors and countries. By construction, cross-country and sector variation in correlated distortions γ_i and entry costs c_e^i match average establishment sizes in all countries. Recall from equation (12) that average establishment size depends on both correlated distortions and entry costs. Since entry costs are derived as a residual, we are interested in the extent to which cross-country variation in γ_i

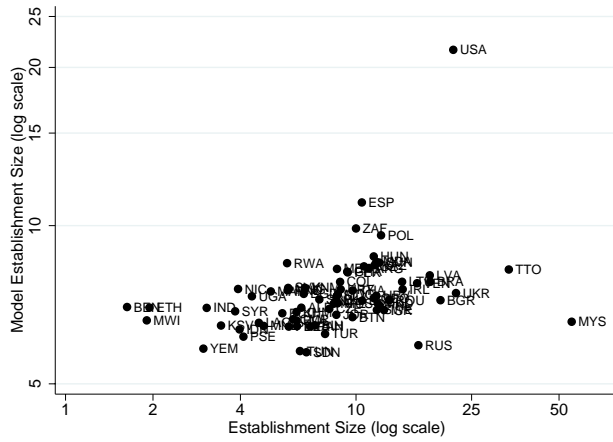
Table 9: MODEL RESULTS ACROSS CORRELATED DISTORTIONS γ

γ	Establishment Size		Endogenous Productivity (Both)	Impact from Misallocation		Sectoral Output	
	(M)	(S)		(M)	(S)	(M)	(S)
0.09	22	5.0	1	1	1	1	1
0.3	14	3.3	0.84	0.96	0.95	0.92	0.92
0.5	11	2.5	0.72	0.88	0.89	0.80	0.80
0.7	9	2.0	0.59	0.78	0.82	0.63	0.67
0.9	7	1.6	0.43	0.67	0.77	0.42	0.48

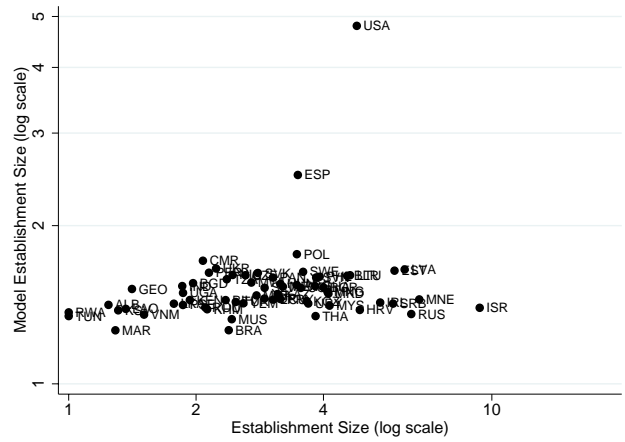
Notes: Results of the model for variations in γ assuming $\gamma = \gamma_m = \gamma_s$. Entry costs and exogenous sector productivity are kept constant. Columns report equilibrium values for average establishment size in each sector, endogenous sectoral productivity, sectoral productivity from factor misallocation, and real sectoral output. Results in columns 4 through 8 are reported relative to the benchmark U.S. economy.

can account for the variation in establishment sizes in the data. Figure 5 reports the average establishment size in the data and the model when only γ_i varies for each sector. We note that differences in γ_i in poorer countries vis a vis the benchmark economy generate large differences in average size both in manufacturing and services. For example, the average γ_i across developing countries in our data is 0.78 in manufacturing and 0.87 in services. This implies an average size of 8 in manufacturing and 1.7 in services. In the data, the average across countries with γ 's close to these averages is an average size of 9 persons engaged in manufacturing and 2.8 in services. These averages are calculated over countries with $\gamma_m \in (0.73, 0.83)$ and $\gamma_s \in (0.82, 0.92)$.

In rationalizing the large difference in average establishment size between the United States and poorer countries in our data, the quantitative model does quite well. At the same time, it is also clear from Figure 5 that the model has a difficult time generating the observed variation in average size among poorer countries. Mechanically, the model predicts a decreasing marginal impact from γ on average size, as documented in Table 9. Given the consistently high γ 's estimated across the (mostly low-income) countries in our sample, the model predicts much less variation in average size across these countries. As a result, the differences in average establishment size predicted by the model due to variation in γ are small compared to the

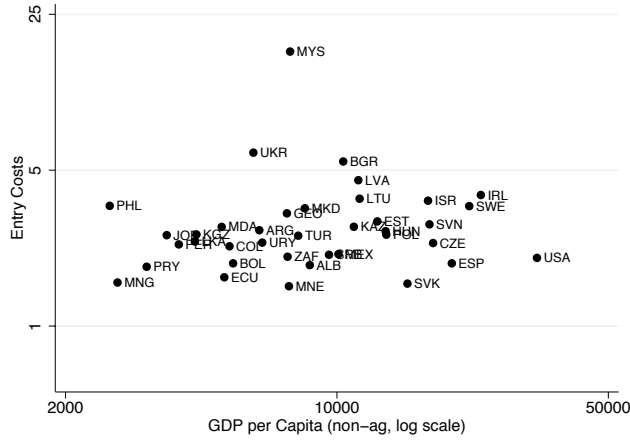


(a) Manufacturing

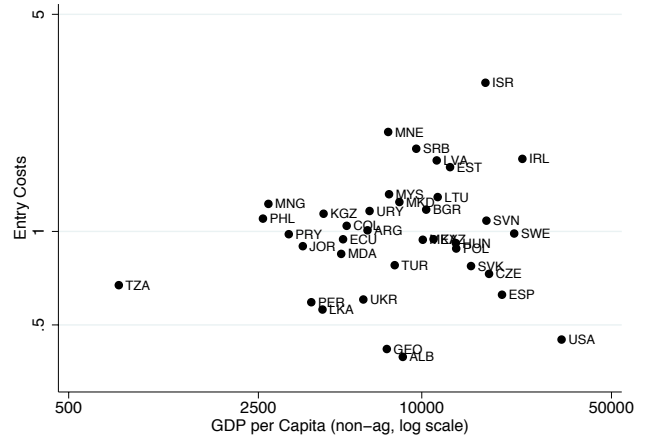


(b) Services

Figure 5: Average Establishment Size, Model vs. Data



(a) Manufacturing



(b) Services

Figure 6: Entry Costs and GDP per Capita

observed size differences. However, the variation in average size unexplained by differences in γ does not appear systematically related to the level of development since the implied entry costs for each country, calculated as a residual to match observed average establishment size in each sector, suggest that entry costs are no higher in poorer countries, if anything they are lower (see Figure 6). While surprising at first sight, this implication of entry costs is consistent with smaller establishments in poor countries.⁸

⁸These results are also consistent with the evidence in [Bento \(2018a\)](#), who uses a model of multi-product firms to decompose observed barriers to competition into product-specific barriers and firm-specific entry costs. He finds entry costs hardly vary across countries and are slightly lower in poorer countries.

To calibrate the model in Section 5.1, we assumed that nonemployers in the U.S. data are establishments with less than three persons engaged (including owner-managers and unpaid workers). If we continue to assume that establishments with less than three persons engaged do not employ paid workers, then we can use the calibrated model to predict how the fraction of nonemployers in a sector should change, given a country’s average establishment size and productivity elasticity of distortions in each sector. While we do not have data related to the size distribution of establishments across countries, we can compare the predictions of the model with data from Sierra Leone, which reports an average establishment size in manufacturing of 2.3 persons engaged, much lower than the average size in the United States (22). We do not have data for the productivity elasticity of distortions for Sierra Leone in manufacturing, but if we assume a value of 0.78 (the average among developing countries), the model predicts that the fraction of establishments that are nonemployers will be 90%, compared to 50% in U.S. manufacturing. This is close to the 83% reported for Sierra Leone’s manufacturing sector, suggesting that average size (and as a consequence, correlated distortions) may be important in accounting for the variation in the share of nonemployer establishments across countries.

Sectoral Productivity The elements affecting sectoral productivity in the model, as specified in equation (13) are: exogenous variation in sector productivity A_i , endogenous entry productivity, establishment size, and misallocation. Figure 7 shows the variation in sector productivity that arises due to variation in γ . As with establishment size, the model generates a substantial drop in sectoral productivity in distorted economies which is about half of that observed in the data. Of the total impact of correlated distortions on sectoral productivity, Table 9 suggests just over half of the impact comes from misallocation across establishments and the rest from the distortionary effects on entry and investment. Relative to the model’s results for differences in average establishment size, Figure 7 shows the model generates variation in sectoral productivity that more closely matches the differences implied by the data.

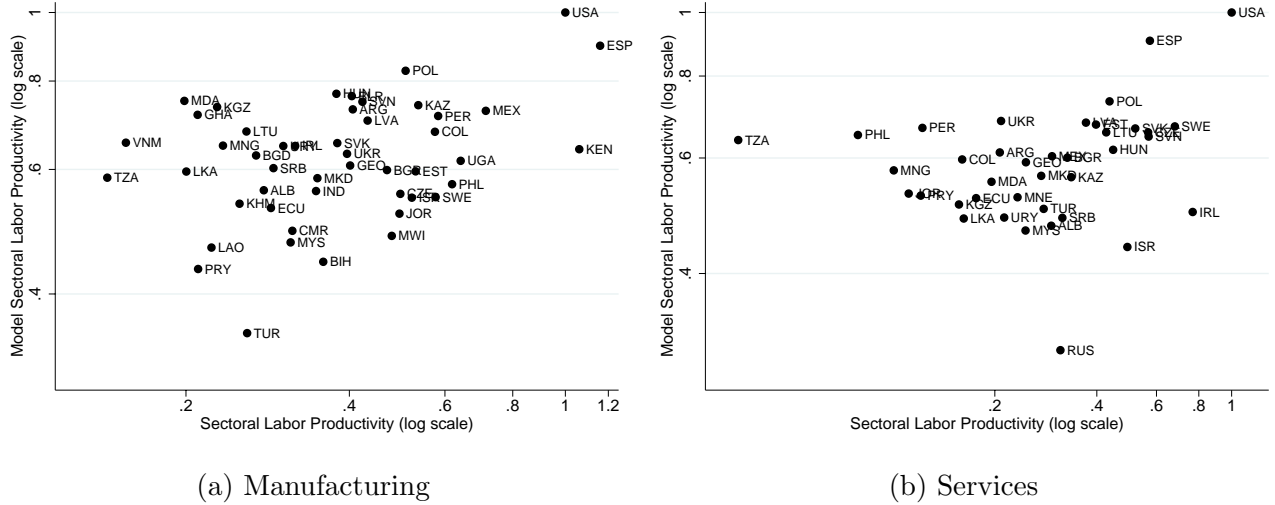


Figure 7: Sectoral Labor Productivity, Model vs. Data

Sectoral Employment and Aggregate Productivity We now explore the implications for sectoral employment and aggregate productivity. Figure 8a illustrates the variation in the employment share of manufacturing relative to services generated by the model with only variation in γ_i across countries and the variation in the data. The model captures about half the variation in the ranges of employment across sectors but we note, as discussed earlier, that there is not a systematic relationship between sectoral shares of manufacturing and services in the non-agricultural sector and GDP per capita in the data. Figure 8b reports the variation in non-agriculture GDP per capita in the model and data. We isolate the effect of differences in observed γ_i across sectors and countries on GDP per capita by applying each economy's γ_i to our benchmark U.S. economy, while keeping A_i and c_e^i constant at U.S. levels in each sector. The model implies substantial variation in non-agricultural GDP per capita and this variation represents almost 50 percent of the variation observed in the data ($\approx \log(3)/\log(10)$).

6 Conclusion

We construct and document comprehensive and comparable data for the average size of establishments across countries within both manufacturing and services. We report a strong positive

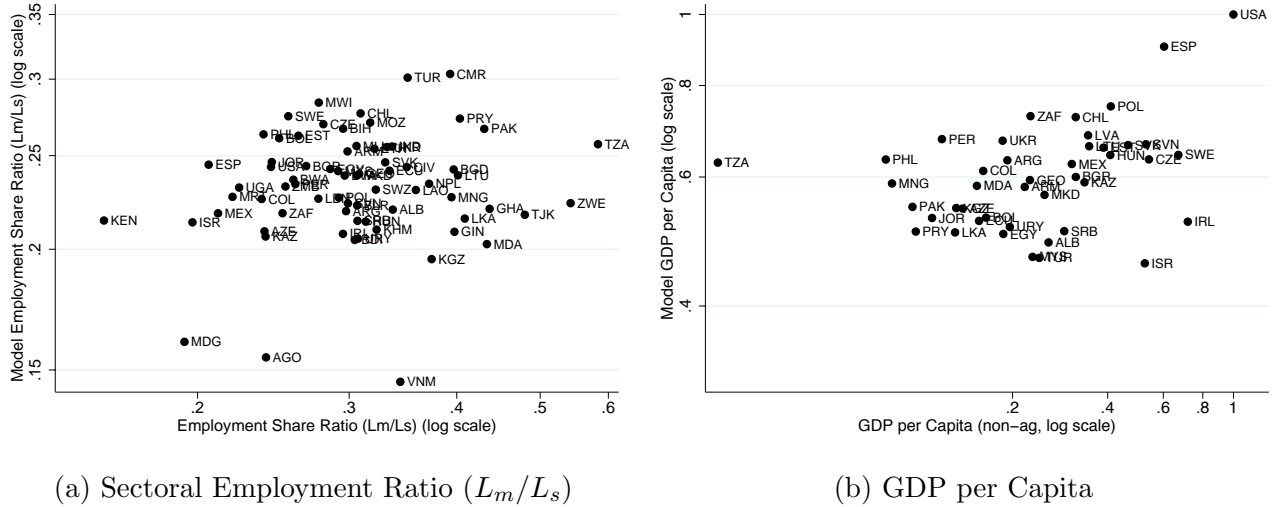


Figure 8: Model versus Data

relationship between establishment sizes and GDP per capita and a remarkably similar income elasticity of average establishment size in manufacturing and services. We also construct and document data for the productivity elasticity of distortions across a large set of countries for manufacturing and service sectors showing a strong negative relationship between this elasticity and GDP per capita in both sectors.

Relating the average size data to several prominent theories of development and TFP, we find that measures of correlated distortions are robustly related to average size and, in particular, to the ratio of average size across sectors. We then consider an otherwise standard model of production heterogeneity with endogenous entry and productivity investment to assess the quantitative impact of policy distortions on establishment size, establishment productivity, and aggregate output. In the theory, entry and productivity investment by establishments is affected by the productivity elasticity of distortions. In a version of the model calibrated to match U.S. data, increasing the productivity elasticity of distortions from 0.09 to 0.7—a level observed for many developing countries in our data—reduces average establishment size in each sector by 60 percent, establishment-level productivity by 41 percent, and output per worker by 37 percent in manufacturing and 33 percent in services.

Our results suggest a strong link between policy distortions (misallocation) and technology

across countries that substantially contributes to differences in establishment size and output per capita across countries. Our evidence on distortions and average establishment sizes across sectors and countries suggest similar impacts of policy distortions on manufacturing and services industries, even though establishment size is generally larger in manufacturing than in services. We also uncover substantial variation in establishment sizes across service industries, which we think is an important issue that warrants further work.

References

- Acs, Zoltan J., Brian Headd, and Hezekiah Agwara**, “Nonemployer Start-up Puzzle,” Technical Report, Small Business Administration, Office of Advocacy 2009.
- Adamopoulos, Tasso and Diego Restuccia**, “The Size Distribution of Farms and International Productivity Differences,” *American Economic Review*, 2014, *104* (6), 1667–1697.
- Atkeson, Andrew and Ariel Tomás Burstein**, “Innovation, Firm Dynamics, and International Trade,” *Journal of Political Economy*, 2010, *118* (3), 433–484.
- Barseghyan, Levon and Riccardo DiCecio**, “Entry costs, industry structure, and cross-country income and TFP differences,” *Journal of Economic Theory*, 2011, *146* (5), 1828–1851.
- Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta**, “Measuring and Analyzing Cross-Country Differences in Firm Dynamics,” in “Producer Dynamics: New Evidence from Micro Data” University of Chicago Press 2009, pp. 15–76.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine**, “Financial institutions and markets across countries and over time: The updated financial development and structure database,” *The World Bank Economic Review*, 2010, *24* (1), 77–92.
- Bento, Pedro**, “Competition as a discovery procedure: Schumpeter meets Hayek in a model of innovation,” *American Economic Journal: Macroeconomics*, 2014, *6* (3), 124–52.
- , “Niche Firms, Mass Markets, and Income Across Countries: Accounting for the Impact of Entry Costs,” *Journal of Economic Dynamics and Control*, 2014, *48*, 147–58.
- , “Competition, Innovation, and the Number of Firms,” Technical Report, Texas A&M University 2018.
- , “Trade Without Scale Effects,” Technical Report, Texas A&M University 2018.
- **and Diego Restuccia**, “Misallocation, establishment size, and productivity,” *American Economic Journal: Macroeconomics*, 2017, *9* (3), 267–303.
- Bernard, Andrew B, Stephen J Redding, and Peter K Schott**, “Multiproduct firms and trade liberalization,” *The Quarterly Journal of Economics*, 2011, *126* (3), 1271–1318.
- Boedo, Hernan J. Moscoso and Toshihiko Mukoyama**, “Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences,” *Journal of Economic Growth*, 2012, *17* (2), 143–170.

- Bollard, Albert, Peter J Klenow, and Huiyu Li**, “Entry Costs Rise with Development,” 2016. Unpublished manuscript, Stanford University.
- Buera, Francisco J, Benjamin Moll, and Yongseok Shin**, “Well-intended policies,” *Review of Economic Dynamics*, 2013, 16 (1), 216–230.
- , **Joseph P Kaboski, and Yongseok Shin**, “Finance and development: A tale of two sectors,” *American Economic Review*, 2011, 101 (5), 1964–2002.
- , – , and – , “Entrepreneurship and financial frictions: A macrodevelopment perspective,” *Annual Review of Economics*, 2015, 7 (1), 409–436.
- Da-Rocha, José-María, Marina Mendes Tavares, and Diego Restuccia**, “Firing costs, misallocation, and aggregate productivity,” Technical Report, National Bureau of Economic Research 2016.
- , – , and – , “Policy distortions and aggregate productivity with endogenous establishment-level productivity,” Technical Report, National Bureau of Economic Research 2017.
- Davis, Steven J., John Haltiwanger, Ronald S. Jarmin, C. J. Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky**, “Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes,” *Producer Dynamics: New Evidence from Micro Data*, 2009, pp. 329–366.
- D’Erasmus, Pablo N and Hernan J Moscoso Boedo**, “Financial structure, informality and development,” *Journal of Monetary Economics*, 2012, 59 (3), 286–302.
- Desmet, Klaus and Stephen L Parente**, “Bigger is better: market size, demand elasticity, and innovation,” *International Economic Review*, 2010, 51 (2), 319–333.
- Dias, Daniel, Carlos Robalo Marques, Christine Richmond et al.**, “Comparing misallocation between sectors in Portugal,” *Banco de Portugal, Economics Research Department Economic Bulletin and Financial Stability Report Article*, 2016, 201602.
- Duarte, Margarida and Diego Restuccia**, “The Role of the Structural Transformation in Aggregate Productivity,” *Quarterly Journal of Economics*, 2010, 125 (1), 129–173.
- and – , “Relative prices and sectoral productivity,” *Journal of the European Economic Association*, 2019.
- Fairlie, Robert W., Javier Miranda, and Nikolas Zolas**, “Job Creation and Survival among Entrepreneurs: Evidence from the Universe of U.S. Startups,” Technical Report, University of California, Santa Cruz 2018.

- Gollin, Douglas**, “Getting Income Shares Right,” *Journal of Political Economy*, 2002, *110*, 458–474.
- Guner, Nezhir, Andrii Parkhomenko, and Gustavo Ventura**, “Managers and productivity differences,” *Review of Economic Dynamics*, 2018, *29*, 256–282.
- Heckman, James and Carmen Pagés**, “Introductory Chapter,” *Law and Employment: Lessons from the Latin America and the Caribbean*, 2004, pp. 1–108.
- Herrendorf, Berthold and Ákos Valentinyi**, “Measuring Factor Income Shares at the Sectoral Level,” *Review of Economic Dynamics*, 2008, *11*, 820–835.
- Herrendorf, Berthold M., Richard Rogerson, and Ákos Valentinyi**, “Growth and Structural Transformation,” *Handbook of Economic Growth*, 2014, *2*, 855–941.
- Hopenhayn, Hugo A.**, “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 1992, *60*, 1127–50.
- Hopenhayn, Hugo A.**, “Firms, Misallocation, and Aggregate Productivity: A Review,” *Annual Review of Economics*, 2014, *6* (1), 735–770.
- , “Firm Size and Development,” *Economía*, 2016, *17* (1), 27–49.
- Hopenhayn, Hugo A. and Richard Rogerson**, “Job Turnover and Policy Evaluation: A General Equilibrium Analysis,” *Journal of Political Economy*, 1993, pp. 915–938.
- Hsieh, Chang-Tai and Peter J. Klenow**, “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.
- and —, “The Life Cycle of Plants in India and Mexico,” *Quarterly Journal of Economics*, August 2014, *129* (3), 1035–1084.
- Kumar, Krishna B., Raghuram G. Rajan, and Luigi Zingales**, “What Determines Firm Size?,” Technical Report, University of Southern California 2001.
- Lucas, Robert E.**, “On the Size Distribution of Business Firms,” *Bell Journal of Economics*, 1978, pp. 508–523.
- Melitz, Marc J.**, “The impact of trade on intra-industry reallocations and aggregate industry productivity,” *Econometrica*, 2003, *71* (6), 1695–1725.
- Melitz, Marc J. and Giancarlo I.P. Ottaviano**, “Market Size, Trade, and Productivity,” *Review of Economic Studies*, 2008, *75* (1), 295–316.

Ordóñez, Julio César Leal, “Tax collection, the informal sector, and productivity,” *Review of Economic Dynamics*, 2014, 17 (2), 262–286.

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

– **and** –, “The causes and costs of misallocation,” *Journal of Economic Perspectives*, 2017, 31 (3), 151–74.

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

Appendix

A Service Sector Establishment Size Data

Table 10 lists each country in the final service sector dataset, the number of industries for which data is available, and the sources from which data have been collected. See [Bento and Restuccia \(2017\)](#) for the corresponding table for manufacturing.

Table 10: LIST OF COUNTRIES AND SOURCES

Country	Code	Industries	Sources
Åland Islands	ALA	5	Statistics and Research Åland: Statistical Yearbook of Åland 2010 and www.asub.ax
Albania	ALB	6	Instituti i Statistikave: www.instat.gov.al/en/figures/statistical-databases.aspx
American Samoa	ASM	9	U.S. Census Bureau: U.S. Economic Census 2007
Andorra	AND	8	Departament d'Estadística: www.estadistica.ad
Anguilla	AIA	8	Anguilla Statistics Department: Abstract of Statistics 2000
Argentina	ARG	8	Instituto Nacional de Estadística y Censos: 2005 Economic Census
Aruba	ABW	8	Central Bureau of Statistics: Business Count 2003
Australia	AUS	9	Australian Bureau of Statistics: Counts of Australian Businesses 2007, Labour Force Surveys (Quarterly)
Austria	AUT	8	Statistik Austria: statcube.at
Bahrain	BHR	8	Kingdom of Bahrain Central Informatics Organization: Population, Housing, Buildings, Establishments and Agriculture Census
Bangladesh	BGD	9	Bangladesh Bureau of Statistics: Economic Census 2013
Belgium	BEL	7	OECD's SDBS Structural Business Statistics
Benin	BEN	9	Institut National de la Statistique et de l'Analyse Economique: General Census of Companies
Bosnia and Herzegovina, Federation of	BIH	5	Institute for Statistics of FB&H: Statistical Yearbook 2012
Brazil	BRA	9	Brazilian Institute of Geography and Statistics: Cadastro Central de Empresas
Brunei	BRN	7	Department of Economic Planning and Development: Brunei Darussalam Statistical Yearbook 2010
Bulgaria	BGR	7	Eurostat
Cambodia	KHM	8	National Institute of Statistics: Economic Census 2011
Cameroon	CMR	5	Institut National de la Statistique du Cameroun: Recensement Général des Entreprises 2009
Canada	CAN	9	Statistics Canada: CANSIM
Cape Verde	CPV	9	Instituto Nacional de Estatística: Estatísticas de Empresas - Inquérito Anual às Empresas 2013

Table 10: LIST OF COUNTRIES AND SOURCES

Country	Code	Industries	Sources
Chad	TCD	8	Institut National de la Statistique, des Etudes Economiques et Démographiques: Recensement Général des Entreprises
Columbia	COL	4	Departamento Administrativo Nacional de Estadística: www.dane.gov.co
Croatia	CRV	8	Eurostat and Central Bureau of Statistics: Statistical Yearbook 2009
Cyprus	CYP	7	Eurostat
Czech Republic	CZE	7	OECD's SDBS Structural Business Statistics
Denmark	DNK	7	OECD's SDBS Structural Business Statistics
Dominican Republic	DNK	7	Oficina Nacional de Estadística: Registro Nacional de Establecimientos
Ecuador	ECU	8	Instituto Nacional Estadística y Censos: National Economic Census 2010
El Salvador	SLV	9	Ministerio de Economica: Tomo I de los VII Censos Económicos Nacionales 2005
Estonia	EST	7	Statistics Estonia: Statistical Yearbook 2011 and pub.stat.ee
Faroe Islands	FRO	7	Statistics Faroe Islands: www.hagstova.fo
Finland	FIN	7	Statistics Finland: www.stat.fi
France	FRA	8	Eurostat
French Guiana	GUF	7	Institut national de la statistique et des études économiques: L'Enquête Annuelle d'Entreprise en Guyane en 2006 and Tableaux Économiques Régionaux Guyane
French Polynesia	PYF	8	Institut de la Statistique de la Polynésie Française: www.ispf.pf and Les entreprises polynésiennes en 2010
FYR Macedonia	MKD	7	State Statistical Office: www.stat.gov.mk
Georgia	GEO	8	National Statistics Office of Georgia: Statistical Yearbooks 2008 and www.geostat.ge
Germany	DEU	7	Eurostat and OECD's SDBS Structural Business Statistics
Ghana	GHA	9	Ghana Statistical Service: Integrated Business Establishment Survey 2014
Greece	GRC	7	Eurostat and OECD's SDBS Structural Business Statistics
Greenland	GRL	7	Statistics Greenland: bank.stat.gl
Guadeloupe	GLP	4	Institut national de la statistique et des études économiques: Caractéristiques des entreprises et établissements and L'Enquête Annuelle d'Entreprise: Les Services en Guadeloupe en 2006
Guam	GUM	9	U.S. Census Bureau: U.S. Economic Census 2007
Guernsey	GGY	9	States of Guernsey: Facts and Figures 2016: Supplementary Data
Hong Kong	HKG	8	Census and Statistics Department: 2007 Annual Surveys of Wholesale, Retail, Import and Export Trades, Restaurants, Hotels, Building, Construction, Real Estate Sectors, Transport and Related Services, Storage, Communication, Banking, Financing, Insurance, and Business Services

Table 10: LIST OF COUNTRIES AND SOURCES

Country	Code	Industries	Sources
Hungary	HUN	7	Eurostat and OECD's SDBS Structural Business Statistics
Iceland	ISL	8	Eurostat
India	IND	8	Central Statistics Office: 2005 Economic Census
Iran	IRN	4	Statistical Centre of Iran: Statistical Yearbook 1389
Ireland	IRL	8	Central Statistics Office: www.cso.ie
Israel	ISR	9	Eurostat and OECD's SDBS Structural Business Statistics
Italy	ITA	7	Eurostat and OECD's SDBS Structural Business Statistics
Japan	JPN	8	Statistics Japan: Establishment and Enterprise Census 2006
Jordan	JOR	8	Department of Statistics: www.dos.gov.jo
Kazakhstan	KAZ	9	Committee on Statistics: www.stat.gov.kz
Kenya	KEN	9	National Bureau of Statistics: Micro, Small and Medium Enterprises (MSMEs) Basic Report 2016 and Statistical Abstract 2016
Korea	KOR	8	Statistics Korea: Censuses on Establishments 2007
Kosovo	KSV	6	Statistical Agency of Kosovo: Statistical Register of Business
Kuwait	KWT	8	Central Statistical Bureau: Annual Surveys of Establishments 2007
Kyrgyzstan	KGZ	9	National Statistical Committee of Kyrgyz Republic: stat.kg
Laos	LAO	9	Lao Statistics Bureau: Economic Census 2006
Latvia	LVA	7	Eurostat
Liechtenstein	LIE	9	Statistical Office: Statistical Yearbooks 2014
Lithuania	LTU	7	Eurostat
Luxembourg	LUX	7	Eurostat
Macau	MAC	5	Statistics and Census Service: Statistical Yearbook 2007
Malawi	MWI	4	National Statistical Office: Statistical Yearbook 2005
Malaysia	MYS	3	Department of Statistics Malaysia: Census of Distributive Trade in 2014
Maldives	MDV	8	Department of National Planning: Economic Survey 2007/2008
Malta	MLT	7	Eurostat
Mauritius	MUS	8	Statistics Mauritius: Censuses of Economic Activity 2002, 2007, Phases I and II
Mexico	MEX	9	Instituto Nacional de Estadística y Geografía: Censos Económicos 2009
Micronesia	FSM	8	Division of Statistics: www.sboc.fm
Moldova	MDA	8	Statistica Moldovei: www.statistica.md
Monaco	MCO	3	Monaco Statistics: Commerce Observatory 2008
Mongolia	MNG	8	National Statistical Office: Mongolian Statistical Yearbook 2011
Montenegro	MNE	8	Statistical Office of Montenegro: www.monstat.org and Statistical Yearbook 2010
Morocco	MAR	9	Haut-Commissariat au Plan du Maroc: 2001-2 Economic Census
Netherlands	NLD	7	Eurostat

Table 10: LIST OF COUNTRIES AND SOURCES

Country	Code	Industries	Sources
New Caledonia	NCL	8	Institut de la Statistique et des Etudes Economique: www.isee.nc
New Zealand	NZL	9	Statistics New Zealand: www.stats.govt.nz
Nicaragua	NIC	9	Instituto Nacional de Información de Desarrollo: Urban Economic Census
Norfolk Island	NFK	2	Australian Business Statistics: www.ausstats.abs.gov.au
Northern Mariana Islands	MNP	9	U.S. Census Bureau: U.S. Economic Census 2007
Norway	NOR	7	Eurostat
Palau	PLW	8	Office of Planning and Statistics: 2012 Economic Indicators
Palestinian Territories	PSE	9	Palestinian Central Bureau of Statistics: Establishment Censuses 2007
Panama	PAN	8	Instituto Nacional de Estadística y Censo: Economic Census 2012
Paraguay	PRY	9	Direccin General de Estadística, Encuestas y Censos: National Economic Census 2011
Peru	PER	9	Instituto Nacional de Estadística e Informática: IV Censo Nacional Economico 2008
Philippines	PHL	9	National Statistics Office: NSO's 2012 List of Establishments
Poland	POL	7	Eurostat
Portugal	PRT	7	Eurostat
Puerto Rico	PRI	9	U.S. Census Bureau: U.S. Economic Census 2007
Qatar	QAT	9	Ministry of Development Planning and Statistics: Establishment Censuses 2008
Romania	ROU	7	National Institute of Statistics: Statistical Yearbooks 2007-2009
Russia	RUS	7	Federal State Statistics Service: Industry of Russia 2009, and Small and Medium Businesses in Russia 2015
Rwanda	RWA	9	National Institute of Statistics of Rwanda: Establishment Census 2011
Samoa	WSM	8	Bureau of Statistics: www.sbs.gov.ws
San Marino	SMR	8	Ufficio Informatica, Tecnologia, Dati e Statistica: Il Bilancio di Previsione per l'Esercizio Finanziario 2010
São Tomé and Príncipe	STP	7	Instituto Nacional de Estatísticas de São Tomé e Príncipe: Business Statistics 2007
Saudi Arabia	SAU	9	Central Department of Statistics and Information: 2010 Economic Census
Serbia	SRB	8	Statistical Office of the Republic of Serbia: Classification Units in the Republic of Serbia 2012, Employees in the Republic of Serbia 2012, and Labor Force Survey 2011
Sierra Leone	SLE	8	Statistics Sierra Leone: Report of the Census of Business Establishments 2005
Singapore	SGP	8	Department of Statistics Singapore: Yearbook of Statistics 2012, 2014, 2015

Table 10: LIST OF COUNTRIES AND SOURCES

Country	Code	Industries	Sources
Slovak Republic	SVK	9	Statistical Office of the Slovak Republic: slovak.statistics.sk and Statistical Yearbook 2013
Slovenia	SVN	7	Eurostat
Spain	ESP	7	Eurostat
Sri Lanka	LKA	7	Department of Census and Statistics - Sri Lanka: Census of Trade and Services 2003-2006
Svalbard	SJM	9	Statistics Norway: www.ssb.no
Sweden	SWE	7	Eurostat
Switzerland	CHE	5	Swiss Statistics: www.bfs.admin.ch/bfs/portal/en/index.html
Taiwan	TWN	9	National Statistics: Industry, Commerce and Service Censuses 2006
Thailand	THA	7	National Statistical Office: Business Trade and Industrial Census 2008 and 2012
Tunisia	TUN	8	Institut National de la Statistique: www.ins.nat.tn
Turkey	TUR	8	OECD's SDBS Structural Business Statistics
Uganda	UGA	8	Uganda Bureau of Statistics: Report on the Census of Business Establishments 2010/2011
Ukraine	UKR	9	State Statistics Service of Ukraine: www.ukrstat.gov.ua
United Arab Emirates	ARE	6	National Bureau of Statistics: www.uaestatistics.gov.ae
United Kingdom	GBR	7	Eurostat
United States	USA	9	U.S. Census Bureau: U.S. Economic Census 2007
Uruguay	URY	8	Instituto Nacional de Estadística: Directory of Companies and Establishments
U.S. Virgin Islands	VIR	9	U.S. Census Bureau: U.S. Economic Census 2007
Venezuela	VEN	8	Instituto Nacional de Estadística: IV Censo Económico
Vietnam	VNM	9	General Statistics Office: Survey of Business Establishments Producing Non-Agricultural Individual Period 2005-2015
Yemen	YEM	8	Central Statistical Organization: Services Survey Report 2004, Transport and Telecom Survey Report 2003, and Internal Trade Survey Results 2004

B Determinants of Establishment Size, Raw Data

We re-estimate the empirical determinants of average establishment size in the service sector and the relative size ratio across sectors using only the raw data. We confirm the main findings using our baseline imputed data. We use the pooled raw size data for service industries, controlling for fixed effects related to both industry and the data used to measure size as described in

Section 2.3. The results of these regressions are reported in Tables 11 and 12. We confirm that all of the estimated coefficients are of the same sign and of similar magnitude.

Table 11: DETERMINANTS OF AVERAGE ESTABLISHMENT SIZE IN SERVICES, RAW DATA

Panel A: Bivariate regressions						
Dependent variable: Average establishment size in services						
Independent variables:						
GDP per capita	0.32***					
	(0.02)					
Services employment		-0.00				
		(0.02)				
External financing			0.29***			
			(0.03)			
Firing costs				-0.19***		
				(0.05)		
Openness to trade					0.34***	
					(0.04)	
Correlated distortions (Serv.)						-1.16***
						(0.18)
Country-Industries	1189	1189	525	525	978	650
R^2	0.40	0.23	0.33	0.21	0.28	0.33

Panel B: Multivariate regressions				
Dependent variable: Average establishment size in services				
Independent variables:				
External financing	0.19***	0.22***		
	(0.04)	(0.04)		
Firing costs	-0.13**		-0.16***	
	(0.06)		(0.05)	
Openness to trade		0.18***	0.14***	0.21***
		(0.06)	(0.05)	(0.04)
Correlated distortions (Mnfg.)				-1.10**
				(0.18)
Country-Industries	388	525	510	620
R^2	0.28	0.35	0.22	0.36

Notes: All variables logged, except for correlated distortions. See the text for the definition of variables and sources. Robust standard errors in parentheses. ***, **, and * refer to one, five, and ten percent levels of significance.

Table 12: DETERMINANTS OF SIZE RATIO MANUFACTURING TO SERVICES, RAW DATA

Dependent variable: Ratio of average size manufacturing to services

Independent variables:

GDP per capita	0.03 (0.03)					
Employment ratio (M/S)	0.03 (0.13)					
External financing		0.16*** (0.04)				
Firing costs			-0.05 (0.05)			
Openness to trade				-0.03 (0.04)		
Correlated distortions gap (M-S)					-1.92*** (0.20)	
Country-Industries	1105	892	496	518	934	622
R^2	0.24	0.27	0.33	0.29	0.28	0.39

Notes: All variables logged, except for correlated distortions. See the text for the definition of variables and sources. Robust standard errors in parentheses. *** refers to a one percent level of significance.