cesifo
WORKING
PAPERS

10447
2023

May 2023

Quantifying the Impact of Red Tape on Investment: A Survey Data Approach

Bruno Pellegrino, Geoffery Zheng

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 10447*

Quantifying the Impact of Red Tape on Investment: A Survey Data Approach

Abstract

An important strand of research in macro-finance investigates which factors impede enterprise investment, and quantifies their aggregate cost. In this paper, we make two contributions to this literature. The first contribution is methodological: we introduce a novel framework to calibrate macroeconomic models with firm-level distortions using enterprise survey micro-data. The core of our innovation is to explicitly model the firms' decisions to report the distortions they face in the survey. Our second contribution is to apply our method across seven countries to characterize the distribution of these distortions and estimate the GDP loss induced by distortionary red tape. Our estimates are based on a dynamic general equilibrium model with heterogeneous firms whose capital investment decisions are distorted by red tape. We find that the aggregate cost of red tape varies widely across the countries in our dataset, with an average cost of of 0.8% of annual GDP. Our framework opens up a new range of applications for enterprise surveys in macro-financial modeling and policy analysis.

JEL-Codes: C830, E200, E600, G380, H100, H200, K200, O100, O400.

Keywords: bureaucracy, growth, investment, legislation, misallocation, red tape, regulations, survey.

Bruno Pellegrino
University of Maryland
College Park / MD / USA
bpellegr@umd.edu

Geoffery Zheng
NYU Shanghai / China
geoff.zheng@nyu.edu

April 2023
Bruno Pellegrino gratefully acknowledges financial support from the Price Center for Entrepreneurship at UCLA Anderson and travel sponsorship from International Association for Applied Econometrics. We are grateful to the editor and an anonymous referee for feedback which significantly improved the paper. We thank Hugo Hopenhayn, Andy Atkeson, Nico Voigtl ander, Simone Lenzu, Max Maksimovic, Ben Moll, Rodrigo Pinto, as well as seminar participants at University of Chicago, UCLA, UC Berkeley, LBS, and the summer Econometric Society/IAAE meetings, for helpful comments and feedback.

1. Introduction

Understanding what keeps markets from allocating resources efficiently is a central research question in economics. A recent literature (Baqaee and Farhi, 2020), inspired by earlier work by Harberger (1954), shows that it is possible to approximate the economic cost of the misallocation of production factors among firms, by measuring the cross-sectional dispersion of firms' markups or revenue productivity (TFPR). It is generally understood that this cost of is large (double-digit percentages of GDP), and that it is higher for emerging economies. Yet, what causes this huge factor misallocation is still very much an open question.

Some studies have shown that information frictions (David, Hopenhayn, and Venkateswaran, 2016), adjustment costs (Asker, Collard-Wexler, and De Loecker, 2014), as well as a variety of country-specific firm size distortions (Gourio and Roys, 2014) are all likely to play a role. However, as shown by David and Venkateswaran (2019), a significant proportion of the variation in revenue productivity and − in particular − in the marginal revenue product of capital (a component of revenue productivity) remains to be explained. This is especially true in the case of emerging economies. Particularly difficult to measure is the impact of institutional factors, such as the quality of the regulatory environment.

In this paper, we seek to make two contributions. The first is methodological. We propose a novel approach to estimate the distortionary impact of *specific* distortions on firm-level investment and factor (mis)allocation. The approach that we propose consists in combining a heterogeneous firms, dynamic general equilibrium model with enterprise survey data. The cornerstone of our methodological contribution is to explicitly model the firms' decision to report the distortions in the survey. This allows us to parametrize the model using moments of the joint distribution of enterprise financials and survey data while also allowing for noisy measurement error in firms' survey responses.

Our second contribution is to then apply our method to estimate the economic cost of red tape (bureaucracy and regulation that delay or constrain investment) across seven developed European economies. We estimate that, across the seven countries in our dataset, the economic cost of red tape exceeds US$ 154 billion each year. When computed at the country level, as a percentage of GDP, the cost of red tape varies widely across countries: it can be as low as 0.10% of GDP, as in the United Kingdom, or as high as 3.9% of GDP, as in France. Interestingly, we find substantial heterogeneity in the economic cost of red tape even among developed European countries.

While this is surely not the first paper to use survey data, and there are many institutions that run firm surveys on a regular basis, this is (to the best of our knowledge) the first paper to embed the firm's survey response decision in a general equilibrium model and to use the survey data to discipline the model.[1] This approach allows us to produce a dollar estimate of the economic cost of red tape for a large set of countries.

A key advantage of our approach is that, unlike previous methods, it allows specifically for measurement error contaminating empirical estimates of marginal revenue products. It also allows for measurement error in the firms' survey responses. Unlike first-generation studies of factor misallocation (Hsieh and Klenow, 2009; Baqaee and Farhi, 2020), in which measurement error in revenue productivity and markups biases upwards the measured deadweight losses, our approach is conservative, and can attribute variation in revenue productivity and markups to specific frictions (Haltiwanger et al., 2018). The tradeoff is that it requires the availability of balance sheet as well as enterprise survey data. Fortunately, this type of dataset is becoming increasingly available to researchers.

[1] In a contemporaneous paper, Ameriks et al. (2016) use investor survey responses to identify preference parameters and estimate a partial equilibrium life-cycle model of insurance demand. Like us, they explicitly model measurement error in the survey responses.

Figure 1: GDP and Capital per Employee v.s. Regulations Index

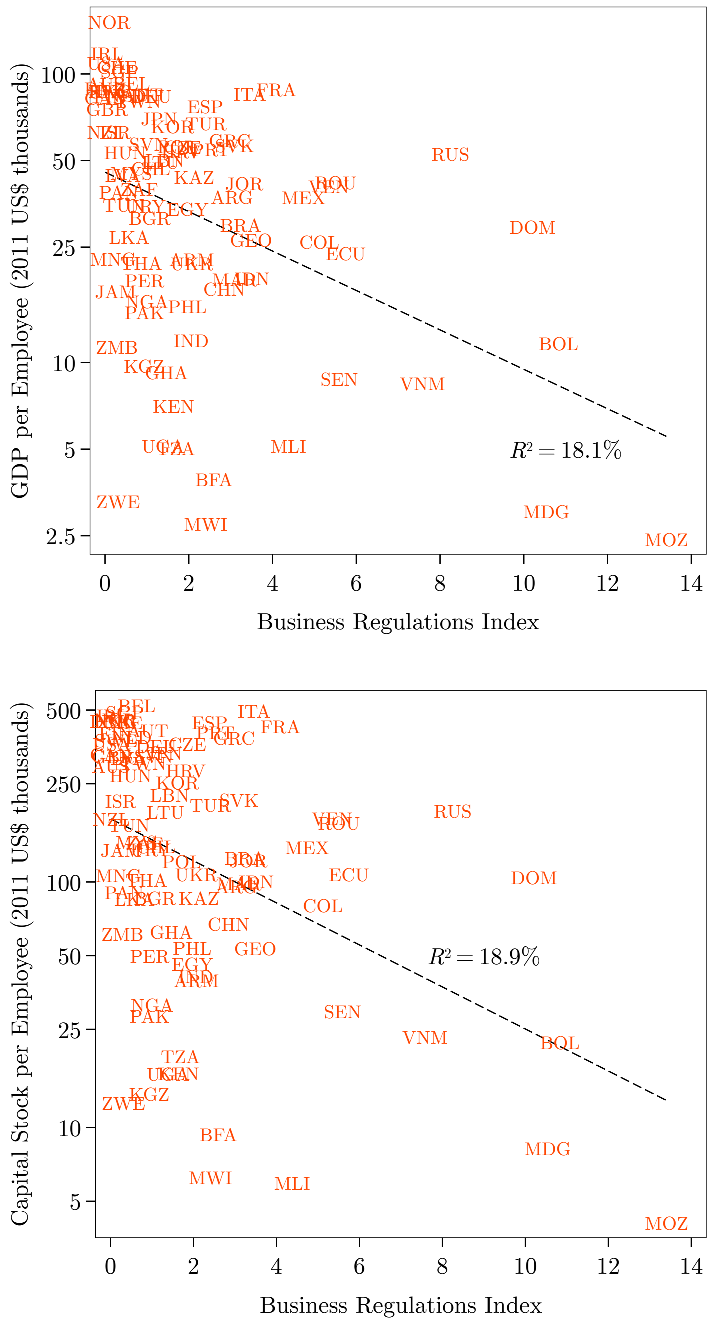

Figure Notes: The figure above plots GDP per employed person (upper panel) and Capital Stock per employed person (lower panel) in 2011 US$ thousands, against an index of business regulations computed from the dataset of Djankov, La Porta, Lopez-de Silanes, and Shleifer (2002). Each observation is a country. The index is the first principal component (in logs) of three variables capturing the average number of days, procedures and steps required to register a new business in the corresponding country. The variables plotted on the vertical axis, which uses a log scale, were obtained from the Penn World Tables v9.1.

We believe that measuring the aggregate economic cost of red tape is an important research question in its own regard. Bureaucracy and regulations have long been argued to cause under-investment and capital misallocation (Posner, 1975; Gray, 1987). Yet, there is a dearth of studies dedicated to measuring aggregate distortionary impact of (broadly defined) red tape on investment.

To motivate our analysis of the relationship between aggregate investment and red tape, we start from some simple empirical observations, for which we use the dataset of business regulations of Djankov, La Porta, Lopez-de Silanes, and Shleifer (2002, henceforth DLLS). DLLS developed a methodology to measure cross-country differences in regulatory burdens and found that countries differ greatly in terms of the red tape faced by their companies.

The upper panel of Figure 1 plots GDP per employee in 2011 US dollars (as measured by the Penn World Tables), against an index of business regulations, which we obtained from the dataset of DLLS. The graph shows a strong negative correlation between these two variables. In the lower panel we replace GDP per employee with capital stock per employee. We find a very similar relationship, except for a much steeper slope – suggesting that the correlation between income and regulation is possibly driven by capital accumulation.

These graphs only present a correlation, of course. What we need in order to estimate the economic cost of red tape is a quantitative model. To derive a precise model-based estimate of the cost of red tape, we propose a dynamic general equilibrium model in which red tape has heterogeneous effects on company investment. We represent the influence of bureaucracy as a capital tax that varies across firms. Our model features endogenous investment, which allows us to disentangle two margins of the adverse effect of bureaucracy on output. The first is under-investment: by imposing a positive tax on capital, red tape discourages investment across all firms. The second channel is capital misallocation: the cross-sectional heterogeneity in the impact of red tape distorts the allocation of capital among firms.

Our model produces closed-form formulas for: i) the percentage loss in GDP induced by red tape, ii) the percentage change in aggregate TFP, and iii) the percentage reduction in the capital stock. The second of these isolates the misallocation channel, whereas the third isolates the under-investment channel. These formulas depend on the parameters of the wedges' cross-sectional distribution.

We then estimate our model using our survey-based approach. The critical assumption underpinning our estimation strategy is that the capital wedges due to red tape follow a known probability distribution. Given our setting, we make the assumption that these wedges follow a normal distribution. We parametrize the distribution using the EFIGE data, a unique dataset that combines firm-level balance sheet data from the Bureau Van Dijk with survey data from GfK. The survey prompts firms to report whether red tape imposes a significant constraint on the company's growth.

The intuition behind our approach is that, to the extent that the survey data correctly identifies firms that are most adversely affected by red tape, we should observe a "shift" in the distribution of the marginal revenue product of capital (MRPK). By effectively measuring the magnitude of this shift, we can parametrize the probability distribution of the wedges. Our model can them inform us how removing the distortions associated with red tape can boost aggregate productivity and investment.

Based on our model we estimate, for each of the countries in the EFIGE dataset, the annual GDP loss induced by red tape, decomposed into GDP losses induced by misallocation and those induced by underinvestment. We find that, for most countries, the aggregate TFP loss is a small share of the overall GDP loss; in other words, most of the GDP losses occur due to lower aggregate saving, rather than capital misallocation.

Our paper brings together different literatures in macroeconomics and finance. First, we contribute from a methodological standpoint to the literatures of financial frictions, under-investment and growth (Beck et al., 2005; Buera et al., 2013; Moll, 2014; Restuccia and Rogerson, 2017) as well as the literature on capital misallocation (Harberger, 1954; Gopinath et al., 2017; David and Venkateswaran, 2019; Baqaee and Farhi, 2020), by showing how red tape can act as a drag on capital investment and allocation and by quantifying this effect.

Second, we contribute to the literature on the effects of regulation and red tape on growth (La Porta et al., 1999; Djankov et al., 2002; Coffey et al., 2020). Our innovation with respect to this literature is to provide a quantitative modeling framework that allows us to perform counterfactual exercises for a wide set of countries. Previous studies (Klapper et al., 2006; Ciccone and Papaioannou, 2007) have shown that red tape lowers aggregate productivity by delaying the entry of the most productive firms and reducing intersectoral factor reallocation. Our study shows that red tape also impacts growth through under-investment and capital misallocation among incumbents. Though our paper focuses on the impact of red tape on firms' investment, our work is also related to studies that have estimated the impact of red tape on labor allocation and labor productivity (Bertrand and Kramarz, 2002; Kaplan et al., 2007; Ebell and Haefke, 2009).

We also contribute to the large literature on cross-country differences in income and institutions (Hall and Jones, 1999; McGrattan and Schmitz Jr, 1999; Barseghyan, 2008; Égert, 2016), by showing how distortions of firms' investment decision and resource misallocation act as important channels that mediate the effect of institutions on growth (Williamson, 2000; Alesina et al., 2005).

This paper also relates to a recent literature in economics and finance that utilizes enterprise surveys (see, for example, Ma et al., 2020; Ameriks et al., 2016) to parametrize macroeconomic models with frictions. We believe that the approach developed in this paper can be used to estimate the effect of other frictions. One way to expand this line of research, which we see as particularly promising, is to design novel enterprise surveys that are built specifically to calibrate general equilibrium models with firm-level frictions. Central Banks and Censuses (who manage business registry data and run enterprise surveys on a recurrent basis) are among the entities that can most easily accomplish this. We believe that having a quantitative framework in place to take advantage of firm-level survey data can provide valuable guidance on how to design enterprise surveys for maximum robustness and statistical power.

Finally, our methodology for incorporating survey evidence into the estimation of economic models is well-suited to take advantage of a growing wealth of survey data in finance. In addition to long-running surveys of consumer expectations (University of Michigan, 2021) and managerial forecasts (Federal Reserve Bank of Richmond, 2021), new initiatives have begun which survey institutional investors (Giglio et al., 2021). Our framework is portable and general enough to be mapped into these settings and address other long-standing questions.

The rest of the paper is organized as follows: in Section 2 we introduce our framework to leverage survey data to measure firm-level distortions; in Section 3 we introduce the data we use and show how we apply this framework to estimate aggregate cost of red tape; in Section 4 we outline the DGE model that underlies our measurements and present estimates obtained from a sample of European manufacturing firms; in Section 5 we present robustness checks and analyze the sensitivity of our results; in Section 6, we conclude.

2. Enterprise Surveys and Distortions: a General Framework

Consider a model economy that is in a static or steady-state equilibrium, with a set of active firms indexed by $i \in \mathcal{I}$. The starting point of our analysis is the profit-maximization problem of a generic firm i, which combines a vector of inputs $\mathbf{x}_i$ to produce output y_i using some increasing concave function $y(\cdot)$:

$$y_i \;=\; y(\mathbf{x}_i) \tag{2.1}$$

Firm i acts as a price-taker in input markets, charges unit price p_i (which may depend on y_i) and faces a multiplicative wedge $\exp(\tau_i)$, which is applied to the expenditure on some generic input X. Hence, the optimality condition of firm i with respect to input X takes the form:

$$\log \mathrm{MRPX}_i \;\stackrel{\text{def}}{=}\; \log \frac{\mathrm{d}\, p_i y_i}{\mathrm{d}\, x_i} \;=\; \log p^X + \tau_i \tag{2.2}$$

where MRPX_i is the Marginal Revenue Product of input X; x_i is the amount of input X utilized by firm i, p^X is the input price, which firm i takes as exogenous; and τ_i is a an actual tax or a shadow tax, which is specific to each firm i and therefore distorts the allocation of input X across firms.

We assume that each firm i draws τ_i from some distribution that has cumulative probability distribution function $F_\tau(\cdot)$, which is defined over a subset of the real line $\mathbb{R}$. By assumption, MRPX_i is decreasing in input X_i. Therefore, ceteris paribus, firms that draw a positive (negative) value of τ acquire less (more) of input X than they would otherwise.

Suppose that we have a fully laid out model for this economy, and that we can write a measure of aggregate output Y (it can also be welfare, or any other statistic of interest for this economy), as a function of a set of model parameters and firm-level observables $\mathcal{O}$, as well as the wedge distribution F_τ, which is not known:

$$Y = Y(\mathcal{O}, F_\tau) \tag{2.3}$$

We are interested in performing counterfactual exercises on F_τ – that is, studying how aggregate output Y responds to changes in the wedge distribution F_τ. In order to perform this exercise, we need to first estimate F_τ. In order to do so, we will require survey data regarding firms' beliefs about τ.

Suppose that an econometrician administers a survey to a sample of firms, asking each firm to quantify the impact of the economic friction τ_i using an ordered categorical variable, which describes the intensity of the distortions experienced by firm i:

$$D_i \;\in\; \{1, 2, ..., m\} \tag{2.4}$$

The problem at hand is to formulate an empirical strategy to recover F_τ, based on some available noisy measure of MRPX_i, as well the survey response data D_i. The first contribution we make in this paper – which is a methodological contribution – is to provide such an empirical strategy, which we base on: 1) modeling the firms' response to the survey; 2) making a parametric assumption on F_τ .

The first step consists of modeling the response of firm i to the survey using a statistical model of ordered choice. We start by constructing a partition $\boldsymbol{\mathfrak{B}}$ of the real line $\mathbb{R}$ into m contiguous intervals:

$$\begin{aligned} \boldsymbol{\mathfrak{B}} \;&=\; \{ \quad \mathfrak{B}_1, \qquad \mathfrak{B}_2, \quad ..., \qquad \mathfrak{B}_m \quad \} \\ &=\; \{(-\infty, T_1]\,,\;\; (T_1, T_2]\,,\;\; ...,\;\; (T_{m-1}, +\infty)\,\} \end{aligned} \tag{2.5}$$

Our key assumption is that firm i's response is characterized by the following rule: firm i returns response

$D_i = j$ if and only if the sum of τ_i and and some reporting error ξ_i falls in the j^{th} element $\mathfrak{B}_j$ of the partition $\mathfrak{B}$ – formally:

$$D_i \equiv \sum_{j=1}^{m} j \cdot \mathbb{I}\left\{\tau_i + \xi_i \in \mathfrak{B}_j\right\} \tag{2.6}$$

the reporting error ξ_i is drawn from some distribution F_ξ, which reflects the inherent noisiness of the survey response. It can originate from a variety of factors, such as the reporting choice being made from a different agent that makes the production decision, or simply incorrect assessment of the severity of the constraint. Intuitively, what we have done is to break the range of possible values of τ_i into a series of "buckets" that correspond to ordered categorical answers that firm i can provide.

The second step is to make a parametric assumption on F_τ – that is, we assume that F_τ has some known functional form and that it is parametrized by (unknown) vector $\boldsymbol{\theta}^\tau \in \mathbb{R}^{\Theta_\tau}$. Similarly, we assume that the survey error is drawn from a distribution with density F_ξ with known functional form and parametrized by unknown $\boldsymbol{\theta}^\xi \in \mathbb{R}^{\Theta_\xi}$.

Additionally, let us define M, the conditional expectation function for τ_i, which conditions on the survey response D_i:

$$M\left(j\,;\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi, \boldsymbol{T}\right) \stackrel{\text{def}}{=} \mathbb{E}\left(\tau_i \middle| D_i = j\right) \tag{2.7}$$

Assume that we observe MRPX with a multiplicative error term, which is assumed to be identically and independently drawn from some distribution with cumulative probability $F_\varepsilon\left(\cdot;\theta^\varepsilon\right)$ that has mean zero, and that it is orthogonal to τ_i and ξ_i, so that:

$$\log \widehat{\text{MRPX}}_i = \log \text{MRPX}_i + \varepsilon_i = \log p^X + \tau_i + \varepsilon_i \qquad \text{with} \qquad \varepsilon_i \sim iid\, F_\varepsilon\left(\cdot;\boldsymbol{\theta}^\varepsilon\right), \mathbb{E}\left(\varepsilon_i \middle| \tau_i, \xi_i = 0\right) \tag{2.8}$$

Define also $\mathcal{P}_j$, the percentage of firms responding $D_i = j$:

$$\mathcal{P}_j \stackrel{\text{def}}{=} \mathbb{P}\left(D_i = j\right) \tag{2.9}$$

as well as $\beta^{(t)}$, the conditional mean difference of $\log \widehat{\text{MRPX}}_i$ between two contiguous response groups:

$$\beta_j \stackrel{\text{def}}{=} \mathbb{E}\left(\log \widehat{\text{MRPX}}_i \middle| D_i = j\right) - \mathbb{E}\left(\log \widehat{\text{MRPX}}_i \middle| D_i = j-1\right) \tag{2.10}$$

The sample counterpart of these two statistics – $\hat{\beta}_j$ and $\hat{\mathcal{P}}_j$ – can measured in the data. Then, the orthogonality assumption $\varepsilon_i \perp \tau_i$ implies that the following system of moment equations:

$$\hat{\beta}_j = M\left(j\,;\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi, \boldsymbol{T}\right) - M\left(j-1;\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi, \boldsymbol{T}\right) \qquad j = 2, ..., m \tag{2.11}$$

$$\hat{\mathcal{P}}_j = F_{\tau+\xi}\left(T_j;\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi\right) - F_{\tau+\xi}\left(T_{j-1};\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi\right) \qquad j = 1, 2, ..., m-1 \tag{2.12}$$

where (with some abuse of notation) $T_0 = -\infty$. This system of moment equations can be used to identify the parameters $\left(\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi, \boldsymbol{T}\right)$. The system above has $2\,(m-1)$ equations and $(\Theta_\tau + \Theta_\xi + m - 1)$ parameters, which include the vector $\boldsymbol{\theta}$ and the $m-1$ thresholds $(T_1, T_2, ..., T_{m-1})$.

In practice, whether $\left(\boldsymbol{\theta}^\tau, \boldsymbol{\theta}^\xi, \boldsymbol{T}\right)$ can be identified depends on the dimensionality of the parameter vectors, the specific functional form of f_τ and f_ξ, and what else can be observed in the data.

3. Application: Red Tape and Capital Investment

In this section, we apply the framework presented in Section 2 to measuring the economic cost of red tape. We start by describing the data, and we then move to the practical implementation of our framework.

3.1. Data

For the practical implementation, our data source is EFIGE: it is a firm-level database that was created by the Brussels-based think tank Bruegel. The dataset covers a representative sample of 14,759 manufacturing firms from seven European countries (Austria, France, Germany, Hungary, Italy, Spain, UK).

The dataset is comprised of two parts. The first is cross-sectional response data from the EFIGE executives survey, which was conducted by the think tank Bruegel in early 2010: firms were asked questions about a wide range of topics, including their organizational structure, ownership, workforce, international activities, and financing. The survey was contracted to the professional marketing research firm GfK.

The second part is a firm-year panel of firm financials (including turnover, assets, interest expenditure, profit and labor costs) for the period 2001-2014 merged from the Amadeus dataset, by the Bureau van Dijk.

We use a dummy variable that encodes the firms' answer to a specific question from the EFIGE survey. The specific question is:

E6. Indicate the main factors preventing the growth of your firm:
- ☐ *financial constraints*
- ☐ *labour market regulations*
- ☐ *legislative or bureaucratic restrictions*
- ☐ *lack of management and/or organizational resources*
- ☐ *lack of demand*
- ☐ *other*

This question admitted an open-ended answer and was coded by the survey administrator as a multiple-choice question. Our main explanatory variable — the dummy variable *Red Tape* — equals one if the survey administrator coded part of the firm's manager answer as indicating option three. That is, the dummy is equal to one if the manger indicates legislative or bureaucratic restrictions were among the main factors preventing their firm's growth. We also encode firms ticking option two as an additional control variable, which we call *Labor Regulations.*

The EFIGE survey asked firms about their activities in 2009. Given that 2009 was a recession year, we use the balance sheet data from the most recent prior expansion year, 2007, for our firm-level analysis.

The survey portion of the EFIGE dataset comes with sampling weights to ensure the representativeness of the survey sample. Weighting ensures that the in-sample distribution of firms over industries and size classes matches the population's.

We model bureaucracy as a shadow tax on capital investment. Hence, our key firm-level variable will be the Average Return on Capital, which in our model is given by y_i/k_i. In the data, we estimate it as follows:

$$\frac{y_i}{k_i} = \frac{\text{Value Added}_i}{\text{Fixed Assets}_i} \tag{3.1}$$

Value Added can be computed as either revenues less intermediate input costs, or (alternatively) as the sum of capital and labor compensation (EBITDA+Labor Costs).

Figure 2: Regulations and Survey Data

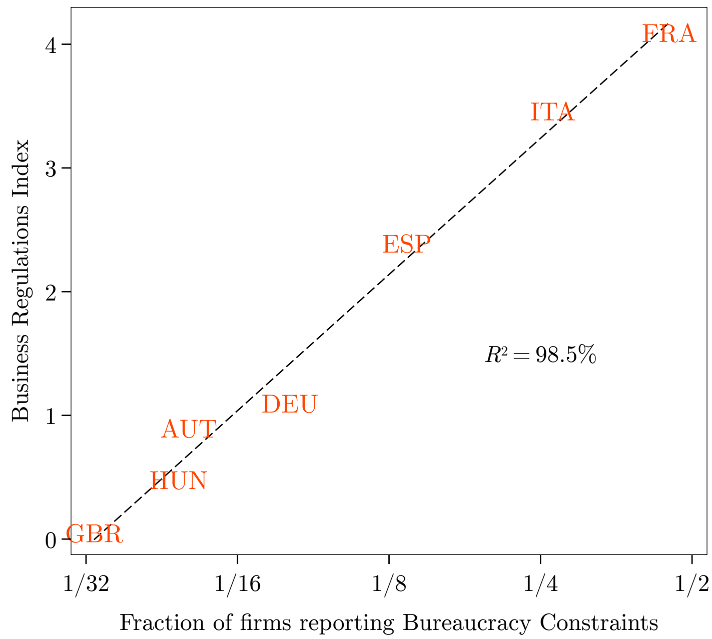

Figure Notes: The figure above plots the probability of a firm reporting bureaucracy as a constraint in the EFIGE survey, by country, against an index of regulatory burden computed from the dataset of Djankov, La Porta, Lopez-de Silanes, and Shleifer (2002). The dotted line is a fitted regression line.

3.2. External Validation of EFIGE Survey Response Data

One risk of using survey data is that the information contained in the *Red Tape* dummy might contain a significant subjective component: in other words, it is difficult to know ex-ante to what extent the responses to the EFIGE survey reflect objective obstacles faced by firms due to red tape, or simply the management's subjective assessment of said obstacles. For this reason, we validate the *Red Tape* dummy variable against a well-known, objective measure of regulatory burden.

In order to validate the EFIGE survey data, we use international data on the intensity of regulations compiled by Djankov, La Porta, Lopez-de Silanes, and Shleifer (2002). The data set covers 85 countries and quantifies the regulatory burden faced by entrepreneurs who intend to form a new business, in terms of number of procedures and the length of time required to complete them.

We use this dataset because it measures red tape across countries with an objective, comparable methodology. The downside is that these measures of regulation are (conceptually) not a perfect match for our model and firm-level data, as neither the model nor the EFIGE survey focus specifically on startups. In order to work with this dataset, we make the explicit assumption that countries that impose more severe constraints on new entrants *also* impose more severe constraints on incumbents. Under this assumption, DLLS's dataset can still serve as a useful proxy. A different way to say this is that we use country-level startup regulations as a proxy for business regulations in a more general sense.

To construct the country-level Regulation Index, we take the first principal component of the logarithms[2] of

[2]Principal component analysis is designed to be done on unbounded variables, whereas the measures in DLLS are positive

the three main measures from this dataset: (i) the number of formal procedures, (ii) the number of steps, and (iii) the average duration in days. The principal component is a natural choice of methodology because it captures common variation across the three factors. In this way, we can be agnostic as to whether Red Tape disproportionately affects specific measures, and instead focus on the cross-country variation in bureaucratic impediments.

Red tape can potentially affect all three of these measures: therefore, we combine them into a single Regulations Index for our study.

Next, we use the Regulations Index to validate EFIGE's survey data. Figure 2 compares, for the 7 countries in EFIGE, the share of firms reporting Bureaucracy as a significant constraint growth, to the corresponding Regulations Index from DLLS's dataset. The number of observations in the plot is small, but the correlation between these two variables is nearly perfect (the R^2 is 98.5%). This is important because it confirms that EFIGE's survey data includes a substantial objective element, which can be linked to a *de jure* measure of regulatory intensity.

3.3. Exploratory Reduced-form Analysis

We begin with a simple exploratory analysis. For our sample of European firms in the EFIGE sample, we show that, consistently with our model, firms which report being constrained by bureaucratic red tape in the EFIGE survey also exhibit higher average Marginal Revenue Product of Capital (MRPK) in the BvD financials database. We compute a first-pass estimate of MRPK as:

$$\text{MRPK}_i \;=\; \frac{\eta - 1}{\eta} \cdot \frac{\partial \log y_i}{\partial \log k_i} \cdot \frac{y_i}{k_i} \tag{3.2}$$

where η is the absolute value of the demand elasticity, which we set to 5, and the elasticity of output with respect to capital is set to 1/3.

Figure 3 plots the raw empirical estimate of the cumulative density function of the log of MRPK, conditional on firms' survey response (constrained/unconstrained). The red dashed line corresponds to those firms that report being unconstrained by red tape ($D_i = 1$), while that plotted as a shaded area corresponds to firms that report being constrained by red tape ($D_i = 2$). In line with our choice to represent bureaucracy as a wedge on the price of capital price, we find that the distribution of log MRPK for constrained firms exhibits first-order stochastic dominance over that of unconstrained firms.

The value of our structural estimation approach is that it enables us to use this observed first-order stochastic dominance in the empirical distribution of $\log \text{MRPK}$ as a way to estimate the distribution of τ. In doing so, we are able to account for the error-in-variables problem that arises due to noisy survey responses as well as additionally control for country- and sector-specific effects using a fixed-effects regression. In the next sub-section, we outline our parametric implementation and our estimation strategy.

3.4. Parametric Implementation

We now show how to implement the methodology that we presented in Section 2. Recall that the objective is to recover the distribution of wedges F_τ. To accomplish this, we impose some additional assumptions: these additional assumptions are informed by the firm-level data we have at our disposal.

by nature of their construction. Thus, we take logarithms of the three measures to make the data amenable to our analysis. The three measures are highly correlated within-country, so that first principal component captures over 98 percent of the cross-country variation in the data.

Figure 3: Conditional Distribution of MRPK

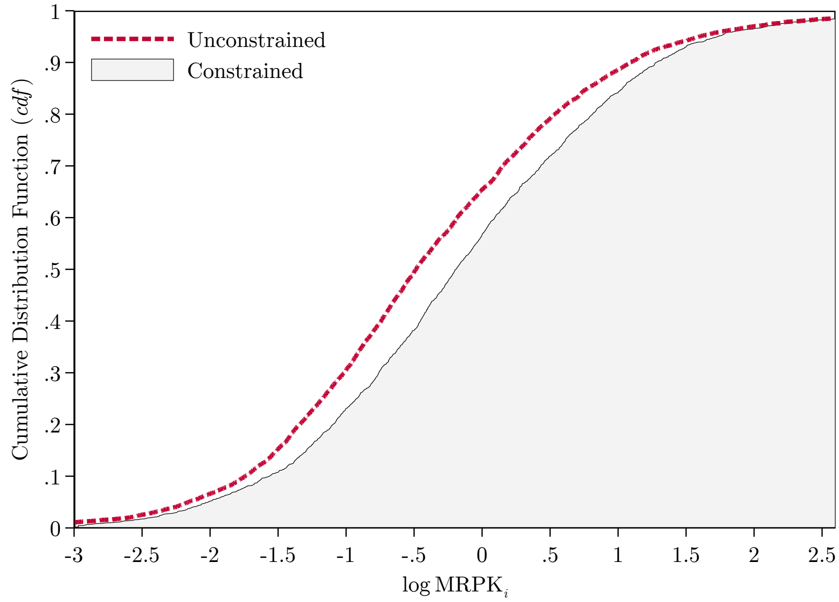

Figure Notes: The figure above plots parametric estimates of the conditional density of the log of Marginal Revenue Product of Capital (MRPK), conditional on the value of the survey dummy in which firms may report bureaucracy as a significant constraint to growth. The grey area is the estimated density for firms that do *not* report bureaucracy as a constraint to growth ($D_i = 1$). The dotted dark line is the estimated density for firms that *do* report bureaucracy as a constraint to growth ($D_i = 2$).

We start from a parametric assumption on F_τ. Given that we are using a firm-level dataset that covers multiple countries, and that we want to estimate the effect of red tape for each individual countries, we assume a functional form for F_τ that is common across all countries, but with a country-specific parametrization.

In what follows, we use the following notation: $\varphi(\cdot)$ is the probability density function of a standard normal, while $\Phi(\cdot)$ is the cumulative distribution function of the standard normal. Also, we use Σ to denote a variance and σ to denote the corresponding standard deviation $\left(\sigma = \sqrt{\Sigma}\right)$.

Assumption 1. *We assume that, for a firm i located in country c, the vector $[\tau_i, \xi_i, \varepsilon_i]$ is drawn from a multivariate normal distribution with country-specific mean vector and diagonal variance-covariance matrix:*

$$\begin{bmatrix} \tau_i \\ \xi_i \\ \varepsilon_i \end{bmatrix} \sim iid\mathcal{N}\left(\begin{bmatrix} \mu_c^\tau \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_c^\tau & 0 & 0 \\ 0 & \Sigma_c^\xi & 0 \\ 0 & 0 & \Sigma_c^\varepsilon \end{bmatrix} \right) \qquad \text{for } i \in c \tag{3.3}$$

This assumption immediately implies that, within each country, the marginal revenue product of capital (MRPK_i) follows a log-normal distribution. Because τ_i is a tax on capital in our empirical application we focus on the dispersion in MRPK. Next, we impose an assumption regarding the measurement error present in the survey responses.

Assumption 2. *We assume that, for all countries c, the reporting error is mean zero. Furthermore, the ratio of the variances of the distortion and the reporting error is given by a constant ρ that is common across* *countries:*

$$\rho \stackrel{\text{def}}{=} \sqrt{\frac{\Sigma_c^\tau}{\Sigma_c^\tau + \Sigma_c^\xi}} \quad \forall\, c \tag{3.4}$$

This assumption implies that the noise-to-signal ratio of the survey response data is constant across countries and that the variance of ξ_i is equal to $\frac{1-\rho^2}{\rho^2}\Sigma_c^\tau$.

Assumption 3. *Each firm i follows the reporting rule*

$$D_i \equiv \sum_{j=1}^{m} j \cdot \mathbb{I}\{\tau_i + \xi_i \in \mathfrak{B}_j\} \tag{3.5}$$

where $\mathfrak{B}_j = (T_{j-1}, T_j]$, for $j = 1, 2, \ldots, m$, and where $\boldsymbol{T}$ is a vector of reporting thresholds that is common across countries, where $T_0 = -\infty$ and $T_m = +\infty$.

The assumption that the vector of reporting thresholds T is the same for firms from different countries is not strictly necessarily to implement our framework: we make it based on a tradeoff. The (significant) advantage of making this assumption is that it allows to pool the data across all countries in one single estimation procedure; this will us sufficient power to estimate the parameter of interest. The question of power is particularly relevant for the EFIGE dataset, in which the number of firms reporting balance sheet data is relatively small for some countries.

The main disadvantage of imposing this assumption is that it assumes away the possibility that there might be any systematic variation in survey reporting error (ξ_i) across countries: in other words, the survey data must be comparable across countries. We believe this is a reasonable restriction for two reasons. First, this was one of the priorities of the architects of the EFIGE survey. Second, as we discussed in Section 3.2, the country-level probability to designate bureaucracy as an impedimedent growth are nearly perfectly correlated with the DLLS regulation index. This reassures us that country-level variation in the EFIGE survey responses reflects objective differences in red tape, rather than systematic reporting error. Of course, in other contexts, owing to either differences in survey methodologies, this assumption may not be as reasonable.

Having defined the firm-level reporting rule, let $\mathcal{P}_{cj}$ be the percentage of firms i in country c that return response j to the survey question $D_i = j$. Also, let $\mathcal{F}_{cj}$ be the corresponding cumulative percentage:

$$\mathcal{P}_{cj} \stackrel{\text{def}}{=} \mathbb{P}(D_i = j | \, i \in c) \qquad \mathcal{F}_{cj} \stackrel{\text{def}}{=} \mathbb{P}(D_i \leq j | \, i \in c) \tag{3.6}$$

Figure 4 outlines this setup visually. Firms's survey responses follow a cutoff rule, but the cutoff is based on the sum of τ, the underlying distortion, and ξ, the uncorrelated survey error. In practice, this means that firms who report being constrained are more-likely to have higher levels of τ, but the strength of the inference depends on the severity of the measurement error, which is captured by ρ.

Assuming that the wedges are normally distributed, we have:

$$\mathcal{P}_{cj} = \Phi\left(\frac{T_j - \mu_c^\tau}{\sqrt{\Sigma_c^\tau + \Sigma_c^\xi}}\right) - \Phi\left(\frac{T_{j-1} - \mu_c^\tau}{\sqrt{\Sigma_c^\tau + \Sigma_c^\xi,}}\right) \quad \text{and} \quad \mathcal{F}_{cj} = \Phi\left(\frac{T_j - \mu_c^\tau}{\sqrt{\Sigma_c^\tau + \Sigma_c^\xi}}\right) \tag{3.7}$$

From the equation above we evince that the observed response frequencies $\mathcal{P}_{cj}$ only inform us about the z-score of the cutoff T_j. To identify the threshold parameter vector $\boldsymbol{T}$, which depends on the distribution parameters μ_c^τ and Σ_c^τ (on which it depends), we require two additional assumptions. First, we use firm i's

Figure 4: The Firms' Survey Response Decision

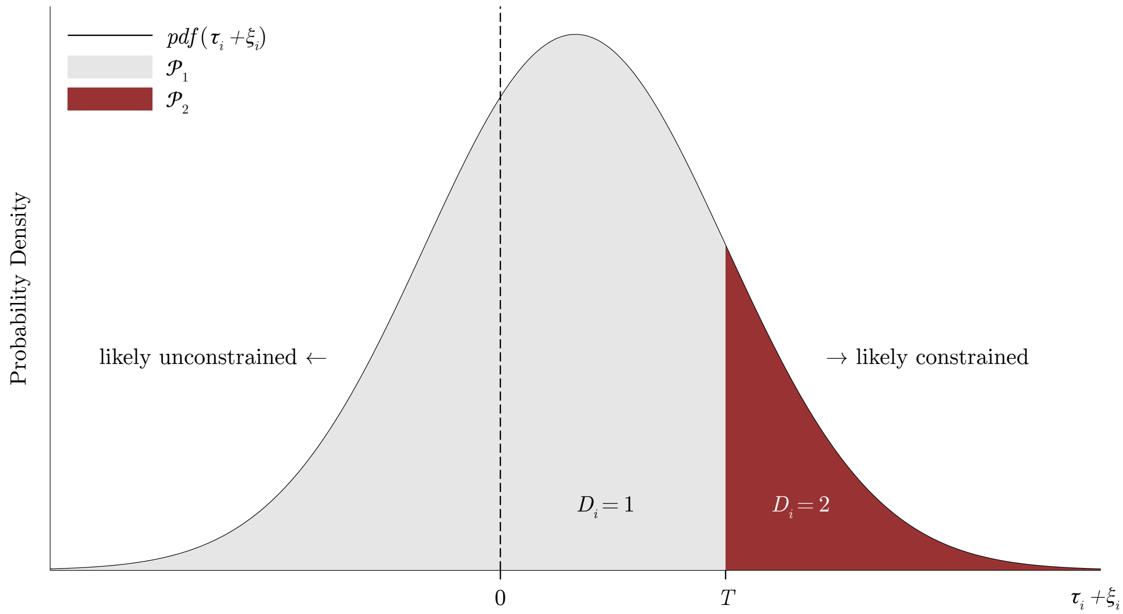

Figure Notes: The above diagram exemplifies firm i's survey response decision.

first order condition for capital, which we augment with an error term to allow for measurement error, in order to capture unobservable variation in MRPK that is unrelated to red tape which we are not modeling explicitly.

Notice that, had we assumed the absence of measurement error ($\varepsilon_i = 0$), the observed dispersion in MRPK would have immediately pinned down the distribution of wedges τ_i. There would not have been any need to use survey data, but we would also have been unable to disentangle red tape from any other source of variation in MRPK (ε). Our approach, which allows measurement error and relies on survey data, allows us to avoid attributing all observed variation in MRPK to capital misallocation.

Under our distributional assumptions, the distribution of $\tau_i + \xi_i$, conditional on country c and survey response $D_i = j$, follows a truncated normal. In the lemma below, we use this fact to derive the conditional mean function of the log MRPK. To this end, it is useful to define the following variable λ_{cj} :

$$\lambda_{cj} \overset{\text{def}}{=} \frac{\varphi\left(\Phi^{-1}\left(\mathcal{F}_{cj}\right)\right) - \varphi\left(\Phi^{-1}\left(\mathcal{F}_{cj-1}\right)\right)}{\mathcal{P}_{cj}} \tag{3.8}$$

λ_{cj} is simply the expected standardized wedge, conditional on $D_i = j$, which we can recover by using the properties of the truncated normal distribution.

Lemma 1. *Given the assumptions above the* (MRPK) *conditional mean function M is given by:*

$$M\left(j; \mu_c^\tau, \Sigma_c^\tau, \Sigma_c^\xi\right) \overset{\text{def}}{=} \mathbb{E}\left(\log \widehat{\text{MPRK}}_i \middle| D_i = j, i \in c\right) = \log r_c + \mu_c^\tau + \rho\, \sigma_c^\tau\, \lambda_{cj}, \tag{3.9}$$

where σ_c^τ denotes the standard deviation of the distribution of τ_i and r is the cost of capital in country c.

Proof. See Appendix A.

□

From Lemma 1, it is clear that comparing differences in average MRPK between survey response groups, which we referred to as β in Section 2, is revealing of σ_c^τ, the country-level dispersion in τ. However, the observed β moments are uninformative regarding the mean of the distortions, μ_c^τ. To infer the mean, we need an additional source of variation in the data. As it turns out, we can use the wording of the survey to identify the mean. To do so, we impose the following assumption.

Assumption 4 (Normalization)**.** *There exists a survey response group j^* for which the expected wedge for firm i, conditional on i reporting response j^* $(D_i = j^*)$, is zero:*

$$\mathbb{E}\left(\tau_i \mid D_i = j^*\right) = 0 \tag{3.10}$$

Assumption 4 effectively requires that some response group j^* is associated with "unconstrained" firms, which is typically the case when we examine enterprise surveys (i.e. there is always an option for firms to indicate that they are unaffected). Given the natural occurrence of such a group, we believe it is a reasonable assumption. Furthermore, our assumption does not constrain the sign of μ_c^τ, and thus is minimally restrictive.

In our setting, we make use of the explicit wording of the EFIGE survey in order to identify the mean-zero response group. Because firms are asked whether bureaucratic restrictions constrain their growth, a natural assumption is that the population of firms which *do not* indicate that they are constrained, the $D_i = 1$ population, has an average wedge of zero.

Finally, given a sector-level Cobb-Douglas production function, we can relate the measured average return to capital to these parameters.

Proposition 1. *Assume that the firm's production function is Cobb-Douglas, with sector (s)-specific output-capital elasticity, and that firm-level demand is isoelastic, with demand elasticity η. Then, the logarithm of the average rate of return on capital satisfies:*

$$\mathbb{E}\left(\log\frac{p_i y_i}{k_i} \mid D_i = j,\, i \in c \cap s\right) = \log\frac{\eta}{\eta-1} + \log r_c + \mu_c^\tau - \left(\frac{\partial \log y}{\partial \log k}\right)_s + T_j \cdot \mathcal{X}_{cj}(\rho) \tag{3.11}$$

where

$$\mathcal{X}_{cj}(\rho) \overset{\text{def}}{=} \frac{\lambda_{cj}}{\frac{1}{\rho^2}\Phi^{-1}\left(\sum_{j'<j}\mathcal{P}_{c,j}\right) - \lambda_{c,1}} \tag{3.12}$$

Proof. See Appendix A. □

Proposition 1 states that the thresholds T_j can be recovered by regressing the (log) average return on capital on $\mathcal{X}_{cj}(\rho)$, a variable which can be computed from the response data by assuming a value for ρ. In the case of the EFIGE survey, because D_i is a dummy variable, there is only one threshold to estimate, and thus only one regression. When D_i is an ordered categorical variable with more than two categories, we must estimate a regression for every threshold.

After estimating the thresholds T_j, we recover the country-specific dispersion of the distortions (Σ_c^τ) using the following expressions:

$$\hat{\sigma}_c^\tau = \frac{\hat{T}_j}{\frac{1}{\rho}\Phi^{-1}\left(\sum_{j'<j}\hat{\mathcal{P}}_{c,j}\right) - \rho\hat{\lambda}_{c,1}} \quad \text{and} \quad \hat{\mu}_c^\tau = -\rho\,\hat{\sigma}_c^\tau\hat{\lambda}_{c,1} \tag{3.13}$$

Figure 5: Validation of Normality Assumption

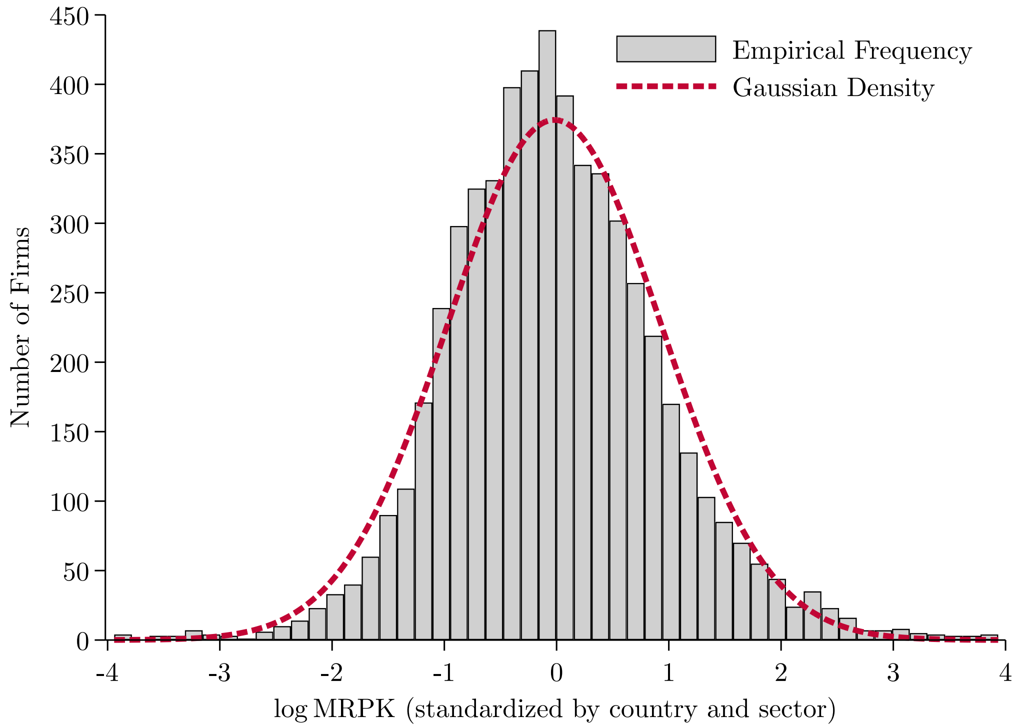

Figure Notes: The figure above plots the distribution of log MRPK, residualized on country and sector fixed effects. A histogram of the data is presented in grey, and the red dashed line is the theoretical density function of the normal distribution. We see that the empirical data is a close match for the normal distribution.

Notice that the estimation of μ^τ and Σ^τ is independent of η and the output-capital elasticity $(\partial y/\partial k)_s$. As we can see from Proposition 1, when we estimate T using a fixed-effect regression, these parameters are absorbed by the country and sector fixed effects. Thus, our econometric estimates of the threshold and of the parameters of F_τ are (by construction) robust to whatever specific values we choose to calibrate these two elasticities.

The plan for our empirical analysis is therefore to : 1) Regress firm-level return on capital $\log \frac{p_i y_i}{k_i}$ on $\mathcal{X}_{cj}(\rho)$ to estimate T; 2) Use $\mathcal{P}_c$ and T to obtain estimates of $[\mu_c^\tau, \Sigma_c^\tau]$. Our estimates of $[\mu_c^\tau, \Sigma_c^\tau]$ will then be used in the next section to parametrize a dynamic general equilibrium model and compute the GDP loss (gain) from red tape. Here, we economize on notation and write expressions containing a single, unique threshold value T. This is harmless when applying our data to the EFIGE dataset, as the firms only indicated whether they were constrained or unconstrained and so there is only one threshold to be estimated.

The calibration of the parameter ρ merits some additional discussion. Unlike a classic errors-in-variable problem, measurement error in the survey response does not introduce an unambiguously downwards bias in our estimates. We first discuss the "classic" situation to develop some intuition. If $\mathcal{X}_{cj}$ were directly observable in our dataset, we would have a typical errors-in-variables correction term in which we can apply a correction to the regression coefficient. In words, because the shift in the distribution of MRPK is observed empirically, a lower chosen value of ρ, corresponding to more noise in the survey responses, would imply a *higher* estimated dispersion. In the case in which the survey is extremely noisy, so that the firms reporting that they are constrained ($D = 2$) is almost identical to the population that reports that they are unconstrained ($D = 1$). In such a setting, if we observe that the constrained firms have a higher average MRPK, it must be that the dispersion in MRPK is sufficiently large that, even just a small change in the conditional distribution of τ results in a large change in the conditional mean of MRPK.

Our situation is complicated by the fact that $\mathcal{X}_{cj}$ itself depends on the calibrated value of ρ, and thus the regression coefficient T_j directly depends non-linearly on the choice of ρ. A priori, the net impact of a more informative survey, corresponding to a higher value of ρ, on the estimated parameters is unclear. Absent clear guidance from the theory, we assess the sensitivity of our results to the chosen value of ρ. We do so in the Appendix by repeating our estimation procedure with different values of ρ.

3.5. Discussion of the Identification Strategy

In deriving Proposition 1, we have implicitly made the assumption that, within a country-sector cell, variation in MRPK reflects a component stemming from bureaucratic restrictions (τ_i) and a component uncorrelated with bureaucratic restrictions (ε_i). This allows us to regress $\log$ MRPK and interpret the regression coefficient as an estimate of the impact of red tape. This assumption was embedded in Equation 2.8, in which we express $\log$ MRPK as the sum of: (i) a market-clearing factor price p^X, (ii) the firm-level distortion arising from red tape τ_i, and (iii) an uncorrelated component ε_i. In this subsection, we discuss a relaxation of this assumption.

An advantage of our regression-based estimation procedure is that we can allow for other determinants of $\log$ MRPK to potentially be correlated with the red tape-induced distortion. To recover an unbiased estimate of the impact of red tape, what we need is that τ_i and ε_i are *conditionally* uncorrelated, given a vector of observables, which we denote $\boldsymbol{v}_i$. Under this relaxed assumption, we need to include these observables $\boldsymbol{v}_i$ in our regression to recover an unbiased estimate of the impact of red tape.

This assumption is inherently data-specific, as it relates to the vector of observable characteristics $\boldsymbol{v}_i$. In our setting, we include the aforementioned country and sector fixed effects, along with controls for the size of the firm, measured in terms of headcount, as well as access to external financing.[3] We report these results in Table 6 in Appendix B.

3.6. Validation of the Normality Assumption

In Figure 5, we validate our normality assumption by plotting the histogram of firm-level $\log$ MRPK, residualized on country and sector fixed effects. Proposition 1 states that, given our underlying assumptions, this variable should be normally distributed. To visually assess this, in Figure 5 we also plot the density function of a normally distributed random variable. We see that, while there is a slight skew right in our data, our empirical histogram of $\log$ MRPK closely conforms to the distribution, making it a suitable choice to use for our analysis both for its convenient mathematical properties as well as its similarity to the distribution we see in the data.

[3] In our model, MRPK and labor demanded are both optimized by the firm taking into account their red tape distortion τ. Thus, a regression controlling for size may underestimate the coefficient of interest due to the so-called "bad control" problem.

4. Model and Quantification

In this section we present a parsimonious dynamic general equilibrium model that incorporates heterogeneous firm-level investment distortions due to red tape. Our model follows broadly David and Venkateswaran (2019), where we extend their analysis by focusing on sources of output losses, rather than TFP losses as they do. By incorporating investment, we arrive at formulas that decompose the overall output loss in our model economy into two distinct channels: misallocation and depressed returns to capital.

4.1. Dynamic General Equilibrium Model

The economy features a representative agent with the following utility function:

$$U\left(C_0, C_1, ...\right) \;=\; \sum_{t=0}^{\infty} e^{-\varrho t} \cdot u\left(C_t\right) \tag{4.1}$$

The function $u(\cdot)$ increasing, concave and twice differentiable. C_t is the consumption of final good at time t and $\exp(-\varrho)$ is the discount rate. The representative agent is endowed with 1 unit of labor which they supply inelastically at a wage rate w_t.

There is a final good producing firm that produces output Y_t, using a CES technology and by taking inputs y_{it} from $i \in [0, 1]$ final good-producing firms:

$$Y_t \;=\; \left(\int_0^1 y_{it}^{\frac{\eta-1}{\eta}} \,\mathrm{d}i\right)^{\frac{\eta}{\eta-1}} \tag{4.2}$$

The final good firm behaves competitively, hence the price of the final good is:

$$P_t \;=\; \left(\int_0^1 p_{it}^{1-\eta} \,\mathrm{d}i\right)^{\frac{1}{1-\eta}} \tag{4.3}$$

and p_i is the price of input i. The intermediate good firms use a Cobb-Douglas production function:

$$y_{it} \;=\; z_i \, k_{it}^{\alpha} \, \ell_{it}^{1-\alpha} \tag{4.4}$$

where k_{it} is capital and ℓ_{it} is labor. In models such as ours, the assumption that distortions τ are statistically independent from productivity z implies that the impact of distortions on aggregate quantities is separable from the distribution of productivity. Therefore, for ease of exposition and to economize on notation, we assume that $z_i = 1$ for all firms. Each firm i rents capital and labor from the representative agent at prices r_t and w_t, respectively. We model the effect of red tape as a firm-specific tax on capital τ_i;

$$\pi_{it} \;=\; p_{it} y_{it} - e^{\tau_i} r_t k_{it} - w_t \ell_{it} \tag{4.5}$$

where τ_i is the shadow tax on capital imposed by bureaucracy/red tape. We assume that both the aggregate profits (Π_t) as well as the aggregate tax bill (G_t) are then rebated back to the consumer:

$$\Pi_t \;\stackrel{\text{def}}{=}\; \int_0^1 \pi_{it} \,\mathrm{d}i \qquad \text{and} \qquad G_t \;\stackrel{\text{def}}{=}\; \int_0^1 \left(e^{\tau_i} - 1\right) r_{it} k_{it} \,\mathrm{d}i \tag{4.6}$$

Let the aggregate capital supply be:

$$K_t \;=\; \int_0^1 k_{it}\,\mathrm{d}i \tag{4.7}$$

Capital depreciates at rate $(1-\delta)$ from period to period, and its evolution is given by the following law of motion:

$$K_{t+1} \;=\; I_t + \delta K_t \tag{4.8}$$

where I_t is aggregate investment, and aggregate consumption is equal to output minus the required investment:

$$\underbrace{P_t Y_t}_{\text{Nominal GDP}} \;=\; P_t\left(C_t + I_t\right) \;=\; \underbrace{r_t K_t + w_t + \Pi_t + G_t}_{\text{Nominal GNI}} \tag{4.9}$$

The consumer's Euler equation is:

$$e^{-\varrho}\, u'\left(C_{t+1}\right)\left(\frac{r_{t+1}}{P_{t+1}} + \delta\right) \;=\; u'\left(C_t\right) \tag{4.10}$$

Demand is isoelastic, therefore firms price at a constant markup over their marginal cost:

$$p_i \;=\; \frac{\eta}{\eta - 1}\, c_i \tag{4.11}$$

where c_i is firm i's marginal cost, including the wedge on capital (τ_i):

$$c_i \;=\; \left(\frac{r e^{\tau_i}}{\alpha}\right)^{\alpha}\left(\frac{w}{1-\alpha}\right)^{1-\alpha} \tag{4.12}$$

The first order conditions for firm i are thus:

$$\text{MRPK}_{it} \;\stackrel{\text{def}}{=}\; \frac{\eta - 1}{\eta}\cdot\alpha\cdot\frac{p_{it} y_{it}}{k_{it}} \;=\; e^{\tau_i} r_t \tag{4.13}$$

$$\text{MRPL}_{it} \;\stackrel{\text{def}}{=}\; \frac{\eta - 1}{\eta}\left(1-\alpha\right)\frac{p_{it} y_{it}}{\ell_{it}} \;=\; w_t \tag{4.14}$$

4.2. Normality Assumption and Aggregate Efficiency

In what follows, we focus on the steady-state equilibrium of this model. The superscript $(^*)$ denotes the relevant model statistic for the counterfactual where the distribution of the wedges degenerates to a mass point at zero (no distortions). In appendix A, we develop expressions for output and productivity for an arbitrary distribution of the wedges τ_i.

We start by defining aggregate steady-state TFP and MRPK (we omit time subscripts to denote the steady state solutions):

$$\text{TFP} \;\stackrel{\text{def}}{=}\; \frac{Y}{K^{\alpha} L^{1-\alpha}} \qquad \text{MRPK} \;\stackrel{\text{def}}{=}\; \alpha\cdot\frac{\eta - 1}{\eta}\cdot\frac{PY}{K} \tag{4.15}$$

In our quantification of the effect of the distortions (and reassured by our previous checks of the normality assumption) we focus on the case in which τ_i is distributed normally with mean μ^{τ} and variance Σ^{τ}. Under this assumption, we can decompose the GDP loss from red tape into two components: one reflects factor misallocation across firms, and can be characterized through the change in aggregate productivity (TFP); the other is related under-investment, and can be characterized through the change in aggregate MRPK.

Figure 6: Estimated Distribution of τ_i

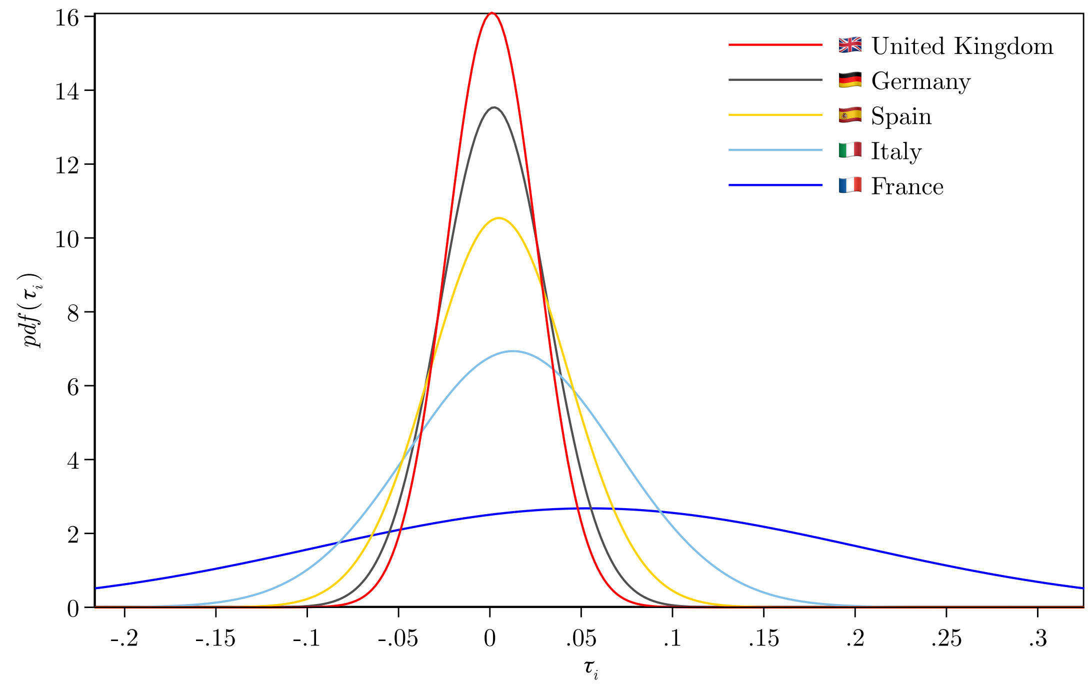

Figure Notes: The figure above plots our estimate of the probability density function of the wedges τ_i, by country. The mean and standard deviation of each normal distribution are obtained via the procedure described in Subsection 3.4.

Proposition 2. *The difference in log GDP between the distorted (observed) steady-state equilibrium and the counterfactual undistorted steady-state equilibrium is:*

$$\log Y - \log Y^* \; = \; \underbrace{\frac{\log \mathrm{TFP} - \log \mathrm{TFP}^*}{1-\alpha}}_{\text{Misallocation}} - \alpha \cdot \underbrace{\frac{\log \mathrm{MRPK} - \log \mathrm{MRPK}^*}{1-\alpha}}_{\text{Under-Investment}} \tag{4.16}$$

where the reduction effect on TFP is given by:

$$\log \mathrm{TFP} - \log \mathrm{TFP}^* \; = \; -\frac{\alpha\,(\alpha\eta + 1 - \alpha)}{2} \cdot \Sigma^\tau \tag{4.17}$$

and the effect on MRPK is given by

$$\log \mathrm{MRPK} - \log \mathrm{MRPK}^* \; = \; \mu^\tau - \left(\frac{2\alpha\eta + 1 - 2\alpha}{2} \cdot \Sigma^\tau \right) \tag{4.18}$$

Proof. See Appendix A. □

From equation (4.17), we see that (leaving firm-level productivity unchanged) aggregate TFP is entirely determined by the dispersion of τ_i. This is well-known from the previous research on capital misallocation. Consistent with economic intuition, the mean of τ_i has no impact on aggregate productivity.[4]

[4] A higher level of μ^τ results in a decline in firm-level factor demand. In equilibrium, this is absorbed by lower factor prices and leaves the cross-sectional distribution of factor inputs, and hence aggregate productivity, unchanged.

Table 1: Estimated parameters of the distribution of τ_i

Country	**Mean** μ^τ **(%)**	**Standard Deviation** σ^τ **(%)**
France	5.45	14.89
Italy	1.27	5.75
Spain	0.50	3.79
Germany	0.23	2.95
Hungary	0.15	2.63
Austria	0.13	2.55
United Kingdom	0.12	2.48

Table Notes: The table above presents our estimates of the mean and the standard deviation of the distribution of τ_i (assumed to be a normal) for each country in the EFIGE dataset. The mean and standard deviation are reported in percentage points and are obtained via the procedure described in Subsection 3.4.

From equation 4.18 we can see that a higher average distortion results in a higher aggregate MRPK. Because capital is dynamically accumulated, an increase in the average distortion is equivalent to a flat tax on the capital stock: it lowers the steady-state capital stock, in turn raising MRPK.

The term in brackets of equation (4.18) is the so-called Oi-Hartman-Abel effect: an increase in the dispersion of τ_i *increases* the steady state capital stock (and thus GDP). This effect is known in the previous literature but more frequently encountered in its application to productivity shocks (Oi, 1961; Hartman, 1972; Abel, 1983): it arises as a consequence of the convexity of firm-level demand for capital. Restuccia and Rogerson (2008), who also have a dynamic model with heterogeneous distortions and endogenous capital, construct their counterfactuals to neutralize this effect.[5] In their setting, as well as ours, the OHA effect can be completely offset by introducing an symmetric tax.

We present our calculation of the GDP losses from red tape both gross and net of the OHA effect. In Subsection 4.4 we present our baseline result, which nets out the OHA effect with an offsetting flat tax as in Restuccia and Rogerson (2008), while in Subsection 5.3 we discuss the results which include the OHA effect.

4.3. Estimated Distribution of the Wedges

We start our quantitative anaysis by presenting the estimated distribution of the wedges. Figure 6 displays the recovered empirical distribution (assumed to be a Gaussian) for each of the countries in our sample.

In Table 1 we tabulate our estimates of the two parameters characterizing the country-specific distribution of τ_i. Given the normality assumption, the mean and the standard deviation are sufficient to fully characterize the distribution.. Proposition 2 states that the misallocation losses are solely a function of the dispersion of τ_i (σ^τ), whereas the under-investment losses due to red tape are driven by high average distortions (μ^τ). For ease of readability, we report parameters in units of percentage points.

In terms of the mean distortion (μ^τ), France exhibits the highest value at 5.45%, implying that, on average, firms in France face the most significant level of wedges due to red tape. Italy comes in second with a mean wedge of 1.27%, followed by Spain at 0.50%, Germany at 0.23%, Hungary at 0.15%, Austria at 0.13%, and

[5] See Footnote 12 of Restuccia and Rogerson (2008), in which they discuss the implications of this "convexity effect."

Table 2: Estimates of the GDP cost of red tape

Country	% Constrained	GDP Loss	Misallocation	Under-Investment
France	45.8%	-3.94%	-1.29%	-2.69%
Italy	26.2%	-0.82%	-0.19%	-0.63%
Spain	14.2%	-0.33%	-0.08%	-0.25%
Germany	7.8%	-0.17%	-0.05%	-0.12%
Hungary	5.4%	-0.12%	-0.04%	-0.08%
Austria	4.8%	-0.10%	-0.04%	-0.07%
United Kingdom	4.3%	-0.10%	-0.04%	-0.06%

Table Notes: The table above presents country-level estimates of the cost of red tape. "% Constrained" is the proportion of that country's firms which report being constrained by bureaucratic regulation. "GDP Loss" denotes the total GDP loss, while "Misallocation" and "Under-Investment" decompose the total loss into the components stemming from misallocation and underinvestment, respectively, using the decomposition presented in Proposition 2. Note that GDP Loss is not exactly the sum of the Misallocation and Under-Investment columns because the effects are additive in logs, not levels.

the United Kingdom at 0.12%. Of the seven countries in the EFIGE dataset, France and Italy have the largest average distortions, while firms in the United Kingdom, Austria, and Hungary generally experience lower levels of wedges compared to their counterparts in other countries.

Regarding the standard deviation (σ^{τ}), we observe the highest value in France (14.89%), implying that there is a substantial degree of variation in the wedges faced by firms in this country. This large variability indicates that some firms in France experience significantly more wedges due to red tape than others. Italy follows with a standard deviation of 5.75%, while the other countries have lower levels of dispersion, with values ranging from 2.48% in the United Kingdom to 3.79% in Spain.

In summary, our analysis indicates that France stands out as the country with the highest average wedges and the most considerable variation in wedges across firms. As we will see, this implies that France has high per-period GDP losses due to red tape. At the other end of the spectrum, the United Kingdom, Austria, and Hungary appear to have lower levels of wedges on average and less variability among firms.

4.4. Quantification of the Aggregate Output Losses

In this section we present our estimates of the GDP cost of red tape, as implied by our model calibrated using our survey-based procedure. Table 2 presents estimates for the 7 countries in the EFIGE dataset.

Following the literature, we calibrate $\eta = 5$ and (based on a Cobb-Douglas production function) $\partial \log y_i / \partial \log k_i = 1/3$ (although, as it will be clear in the empirical section of the paper, these values have no influence our estimate of the wedge distribution).

Consistent with our observation that the mean of the distribution of τ_i is close to zero for the United Kingdom and Germany, we see that the GDP loss due to underinvestment is close to zero. France, which had a very large average τ_i, is estimated to lose roughly 4% per year due to bureaucratic red tape, with 2.7% of that arising from underinvestment and the remaining 1.3% arising from misallocation. We refer to the loss attributed to $\Delta \log \text{MRPK}$ as "underinvestment" because these losses stem from a high average τ_i, which acts as a country-wide tax on capital that depresses investment.[6]

A key advantage of estimating the distribution parameters using our methodology is that it lets us distinguish between: i) countries in which a large fraction of firms report being constrained, and ii) countries in which the distribution of distortions is highly dispersed. In our sample, 26% of Italian firms reported being constrained, but we find a relatively small impact on GDP of 0.8% compared to France, in which 46% of firms report being constrained, and which has a five-fold greater GDP loss of almost 4%. This highlights the importance of a structural approach to the estimation that explicitly models the survey response decision of firms.

Furthermore, we see that geographically and culturally similar countries suffer disproportionately from red tape. We see prima facie evidence of this in the fact that firms in neighboring countries report markedly different levels of constrained-ness. These differences are present in the raw data and suggest that countries suffer disproportionately from red tape. When we apply our methodology, we find that this is indeed the case. Even within a sample of developed European countries, there is substantial heterogeneity in the severity of the GDP losses. We see that France has relatively high GDP losses, whereas neighboring countries Spain and the United Kingdom have GDP losses of 0.3% and 0.1%, respectively.

The prior literature on misallocation had focused on the impact of distortions on a reduction in TFP. Our results highlight that, while misallocation is an important channel through which regulatory burdens handicap countries' economies, it is not the only channel. In general, we find that the GDP impact of underinvestment exceeds that of misallocation by a factor of 2 to 3, and that the total impact of red tape is non-negligible. By way of comparison, Rubbo (2023) estimates the welfare costs of monetary policy and estimates the costs to US consumers to be 0.8% of GDP. Our estimates of the costs of red tape are of similar magnitude on average, and imply that the seven EFIGE countries lose a combined $154.5 billion each year due to red tape, with roughly two thirds stemming from under-investment.[7] Policy improvements can offer major welfare improvements, emphasizing the importance of work understanding and quantifying the consequences of onerous bureaucratic procedures.

[6]We acknowledge that increasing TFP also results in additional investment, as in David and Venkateswaran (2019).

[7]This estimate is based on applying our per-period GDP loss estimates to 2019 Real GDP. GDP data is from the Penn World Tables and is expressed in terms of 2017 USD.

5. Robustness Checks and Further Discussion

5.1. Sample Selection

In this section, we discuss the robustness of our empirical results, and we provide additional discussion of a number of empirical issues. The first issue we seek to discuss in more depth is the construction of our estimation sample.

While the weights in the EFIGE dataset guarantee representativeness of the survey portion of the dataset, there are some well-known issues of coverage and sample selection which affect the BvD Amadeus database (Kalemli-Ozcan et al., 2015), the source of the balance sheet portion of EFIGE. Specifically, firm financials appear to be missing, for certain countries (Austria, Germany, the UK) in a non-random way.

We made sure that our results are not driven by this sample selection into the BvD database. Specifically, we use the fact that France, Hungary, Italy and Spain are virtually free from the sample selection problem (coverage in these countries is nearly 100%). Our baseline estimates also hold when we estimate the threshold T by excluding Austria, Germany and the UK from the Sample. Once we have T, we can still perform model calculations for these countries, since all that the survey portion ($\mathcal{P}_c$) is all we need.

5.2. Parameter Sensitivity Analysis

In our estimation procedure, we have assumed a value of ρ, the paramter which controls the signal-to-noise ratio of the survey, was constant across all countries in the sample, of 0.5: this value implies that the survey responses are equal parts signal and noise.

In this section, we present results for alternative calibrations of ρ. We find that our results are not particularly sensitive to the choice of ρ – that is, that the estimated GDP losses from red tape are of similar economic magnitude.

In the Appendix, Table 3 we repeat our analysis, where ρ is calibrated to be 1/3 and 2/3. The upper panel presents the GDP Loss and its decomposition for ρ equal to 1/3. This value corresponds to an assumption that the variance in survey data due to the true distortions represents one-third of the total variance. For this value of ρ, we tend to estimate overall *larger* GDP losses. This is the case for France, Italy, and Spain. For Germany, Hungary, Austria, and the United Kingdom, we see that the GDP losses are slightly smaller given this value of ρ. Even though the Misallocation effect estimates are unambiguously larger for smaller values of ρ, the change in the Under-Investment effect is ambiguous, leading to the decline in estimated GDP Losses.

The lower panel reports the same set of results for $\rho = 2/3$: this value implies that two-thirds of the variation in the firm's survey response is attributable to the true underlying distortions τ_i; the survey is thus relatively informative about firm's constraints. We see that the estimated GDP Losses are lower than in our baseline calibration for France, but larger for the other countries in our sample.

5.3. The Oi-Hartman-Abel Effect

In this subsection, we provide some more background on the Oi-Hartman-Abel (OHA) effect and how it affects our computation of the GDP cost of red tape. As we have previously shown in section 4.4, the OHA effect consists in a positive impact of an increase in the variance of the distortions (Σ^τ) on the steady-state level of capital.

The economic intuition behind this result lies in the fact that firms endogenously choose their capital stock to avoid distortions and take advantage of subsidies. Firms that draw a large value of τ will endogenously respond by reducing their capital factor demand, while firms which draw a small value of τ will endogenously scale up and rent more capital. Thus, the τ distortions induce convexity in firm-level capital demand and act like a firm-specific real option for each firm in the economy, which acts like a subsidy on capital.

One way to see this effect quantitatively is to compute the average effective tax rate (ETR) on the capital stock, which we define as follows:

$$e^{\text{ETR}} \quad \stackrel{\text{def}}{=} \quad \frac{\int_0^1 e^{\tau_i} r k_i \, \mathrm{d}i}{\int_0^1 r k_i \, \mathrm{d}i} \quad = \quad \int_0^1 e^{\tau_i} \frac{k_i}{K} \, \mathrm{d}i. \tag{5.1}$$

We can use the fact that $K \stackrel{\text{def}}{=} \int_0^1 k_i \mathrm{d}i$ to obtain:

$$e^{\text{ETR}} \quad = \quad \int_0^1 e^{\tau_i} \mathrm{d}i + \operatorname{cov}\left(e^{\tau_i}, \frac{k_i}{K}\right) \quad = \quad \exp\left(\mu^\tau + \frac{1}{2}\Sigma^\tau\right) + \operatorname{cov}\left(e^{\tau_i}, \frac{k_i}{K}\right) \tag{5.2}$$

where, under our assumption that productivity and distortions are uncorrelated, the covariance of firm-level capital demand k_i and the distortion τ_i is unambiguously negative. Therefore, for any positive amount of dispersion in firm level distortions τ_i, the effective tax rate is strictly less than the average distortion (notice that the average distortion includes the Jensen's inequality term). In particular, in the case where the average tax on capital is zero, the presence of distortions will result in a subsidy to capital, shifting the aggregate demand of capital to the right.

In Appendix Table 4, we present results for the decline in output, producitvity, and investment in which we do not offset OHA effect, so that the impact of under-investment is given by Equation (4.18). Using this approach, the impact of under-investment is significantly reduced in relative percents, albeit still present and economically significant.

5.4. Bureaucracy as a Tax on Labor

While our theoretical framework is sufficiently general to account for distortions to various factor inputs, in our empirical application we have focused on an economy in which the distortions due to red tape load on capital. While we see this as the most natural way to model bureaucracy, it is also plausible that red tape might affect labor, and thus lead constrained firms to have higher MRPL.

In the Appendix, Table 6, we investigate this possibility. In column (1), we regress the log of MRPL on D_i (the firm's survey response) along with country and sector fixed effects. We see that the regression slope coefficient is statistically indistinguishable from zero, and if anything constrained firms tend to have lower MRPL. In column (2), we replace the survey indicator with $\mathcal{X}$, as defined in Proposition 1. Unlike the regression slope coefficient for log of MRPK on χ is positive and statistically different from zero, we find that the corresponding one for MRPL is negative and not statistically distinguishable from zero.

Motivated by the lack of an empirical relationship between MRPL and firms' survey responses, we focus on MRPK in our empirical application and leave further investigation of a possible relationship between MRPL and bureaucracy to future research.

6. Conclusions

In this paper, we have proposed a novel measurement framework for micro frictions in macro-financial models that leverages enterprise survey data. Our key innovation is to explicitly model the firm's survey response.

We have used this framework to calibrate a dynamic general equilibrium model with heterogeneous firms, which was then used to study the distortionary impact of bureaucracy and regulation on capital investment and (mis)allocation. We modeled regulations as firm-specific taxes on capital, and computed the aggregate impact of these taxes on both output as well as total factor productivity. We have then taken our model to the data using linked survey and balance sheet enterprise microdata. This enabled us to identify which firms are relatively constrained, as well as observe the impact of these constraints on their performance.

Using a dataset of matched firm-level financials and survey responses covering seven large European countries, we estimated the GDP loss due to red tape. We found that red tape leads to an average annual GDP loss of 0.8% of GDP, but with highly-asymmetric effects across countries: in some countries such as France, red tape adversely and/or unequally affects a large share of the firms, resulting in a loss in value added that approaches 4% of the yearly GDP; in other countries (such as UK or Austria) the effects of red tape are relatively mild and/or uniform across firms, resulting in a loss of a few basis points. Our model allows us to decompose the loss in GDP into two distinct channels: (i) the effects of depressed capital investment and (ii) the effects of an inefficient allocation of resources. We find that both effects are economically significant in magnitude.

Our approach for estimating aggregate effects based on micro-data is fairly general and can be applied across a wide sample of surveys. In addition, we believe that our framework can inform the design of future enterprise surveys. Our results highlight the costs of over-regulation, as well as the importance of institutions for growth. These costs are similar in magnitude to those estimated for business cycles and sub-optimal monetary policy.

As a final remark, we have used theory and survey data to quantify the effects of a specific friction – namely, bureaucratic restrictions on capital. Much variation in marginal revenue products of capital and labor remains to be explained, and we suggest that future work could extend our analysis to new data, in order to address the impact of a wider range of constraints on the economy.

References

ABEL, A. B. (1983): "Optimal investment under uncertainty," *The American Economic Review*, 73, 228–233.

ALESINA, A., S. ARDAGNA, G. NICOLETTI, AND F. SCHIANTARELLI (2005): "Regulation and investment," *Journal of the European Economic Association*, 3, 791–825.

AMERIKS, J., J. BRIGGS, A. CAPLIN, M. D. SHAPIRO, AND C. TONETTI (2016): "The long-term-care insurance puzzle: Modeling and measurement," Tech. rep., National Bureau of Economic Research.

ASKER, J., A. COLLARD-WEXLER, AND J. DE LOECKER (2014): "Dynamic inputs and resource (mis) allocation," *Journal of Political Economy*, 122, 1013–1063.

BAQAEE, D. R. AND E. FARHI (2020): "Productivity and misallocation in general equilibrium," *The Quarterly Journal of Economics*, 135, 105–163.

BARSEGHYAN, L. (2008): "Entry costs and cross-country differences in productivity and output," *Journal of Economic Growth*, 13, 145–167.

BECK, T., A. DEMIRGÜÇ-KUNT, AND V. MAKSIMOVIC (2005): "Financial and legal constraints to growth: does firm size matter?" *The journal of finance*, 60, 137–177.

BERTRAND, M. AND F. KRAMARZ (2002): "Does entry regulation hinder job creation? Evidence from the French retail industry," *the quarterly journal of economics*, 117, 1369–1413.

BUERA, F. J., B. MOLL, AND Y. SHIN (2013): "Well-intended policies," *Review of Economic Dynamics*, 16, 216 – 230, special issue: Misallocation and Productivity.

CICCONE, A. AND E. PAPAIOANNOU (2007): "Red tape and delayed entry," *Journal of the European Economic Association*, 5, 444–458.

COFFEY, B., P. A. MCLAUGHLIN, AND P. PERETTO (2020): "The cumulative cost of regulations," *Review of Economic Dynamics*.

DAVID, J. M., H. A. HOPENHAYN, AND V. VENKATESWARAN (2016): "Information, Misallocation and Aggregate Productivity*," *The Quarterly Journal of Economics*.

DAVID, J. M. AND V. VENKATESWARAN (2019): "The sources of capital misallocation," *American Economic Review*, 109, 2531–67.

DJANKOV, S., R. LA PORTA, F. LOPEZ-DE SILANES, AND A. SHLEIFER (2002): "The regulation of entry," *The Quarterly Journal of Economics*, 117, 1–37.

EBELL, M. AND C. HAEFKE (2009): "Product market deregulation and the US employment miracle," *Review of Economic dynamics*, 12, 479–504.

ÉGERT, B. (2016): "Regulation, institutions, and productivity: new macroeconomic evidence from OECD countries," *American Economic Review*, 106, 109–13.

FEDERAL RESERVE BANK OF RICHMOND (2021): "The CFO Survey," Accessed: 2021-07-29.

GIGLIO, S., M. MAGGIORI, J. STROEBEL, AND S. UTKUS (2021): "Five Facts about Beliefs and Portfolios," *American Economic Review*, 111, 1481–1522.

GOPINATH, G., Ş. KALEMLI-ÖZCAN, L. KARABARBOUNIS, AND C. VILLEGAS-SANCHEZ (2017): "Capital allocation and productivity in South Europe," *The Quarterly Journal of Economics*, qjx024.

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

Gray, W. B. (1987): "The cost of regulation: OSHA, EPA and the productivity slowdown," *The American Economic Review*, 77, 998–1006.

Hall, R. E. and C. I. Jones (1999): "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, 83–116.

Haltiwanger, J., R. Kulick, and C. Syverson (2018): "Misallocation measures: The distortion that ate the residual," Tech. rep., National Bureau of Economic Research.

Harberger, A. C. (1954): "Monopoly and Resource Allocation," *The American Economic Review*, 77–87.

Hartman, R. (1972): "The effects of price and cost uncertainty on investment," *Journal of economic theory*, 5, 258–266.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2015): "How to construct nationally representative firm level data from the ORBIS global database," Tech. rep., National Bureau of Economic Research.

Kaplan, D. S., E. Piedra, and E. Seira (2007): *Entry regulation and business start-ups: Evidence from Mexico*, The World Bank.

Klapper, L., L. Laeven, and R. Rajan (2006): "Entry regulation as a barrier to entrepreneurship," *Journal of financial economics*, 82, 591–629.

La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny (1999): "The quality of government," *The Journal of Law, Economics, and Organization*, 15, 222–279.

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020): "A quantitative analysis of distortions in managerial forecasts," Tech. rep., National Bureau of Economic Research.

McGrattan, E. R. and J. A. Schmitz Jr (1999): "Explaining cross-country income differences," *Handbook of macroeconomics*, 1, 669–737.

Moll, B. (2014): "Productivity losses from financial frictions: can self-financing undo capital misallocation?" *The American Economic Review*, 104, 3186–3221.

Oi, W. Y. (1961): "The desirability of price instability under perfect competition," *Econometrica: journal of the Econometric Society*, 58–64.

Posner, R. A. (1975): "The Social Costs of Monopoly and Regulation," *The Journal of Political Economy*, 807–827.

Restuccia, D. and R. Rogerson (2008): "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic dynamics*, 11, 707–720.

——— (2017): "The Causes and Costs of Misallocation," *Journal of Economic Perspectives*, 31, 151–74.

Rubbo, E. (2023): "Networks, Phillips Curves, and Monetary Policy," *Econometrica (forthcoming).*

University of Michigan (2021): "University of Michigan: Consumer Sentiment [UMCSENT]," Accessed: 2021-07-29.

Williamson, O. E. (2000): "The New Institutional Economics: Taking Stock, Looking Ahead," *Journal of Economic Literature*, 38, 595–613.

Appendix
Quantifying The Impact of Red Tape on Investment: A Survey Data Approach

Bruno Pellegrino – Geoffery Zheng

A. Proofs

A.1. Proof of Lemma 1

We are interested in the condition expectation of τ for $\tau + \xi \in (\alpha, \beta)$,

$$\mathbb{E}\left[\tau \mid \tau + \xi \in (\alpha, \beta)\right]. \tag{A.1}$$

First, define standard normal $U \overset{\text{def}}{=} \frac{\tau - \mu^\tau}{\sqrt{\Sigma^\tau}}$ so that we can rewrite the condition expectation of interest as

$$\mu^\tau + \sqrt{\Sigma^\tau}\mathbb{E}\left[U \mid \alpha < \mu^\tau + \sqrt{\Sigma^\tau}U + V < \beta\right]. \tag{A.2}$$

This new modified conditional expectation can be expressed as

$$\mathbb{E}\left[U \mid \alpha < \mu^\tau + \sqrt{\Sigma^\tau}U + V < \beta\right] = \frac{\int_{\mathbb{R}^2} v\, f_{V\xi}\left(v, \xi\right) \mathbf{1}_{(\alpha-\mu^\tau, \beta-\mu^\tau)}\left(\sqrt{\mathbf{\Sigma}^\tau}\mathbf{U} + \mathbf{V}\right) \mathrm{d}v\, \mathrm{d}y}{\int_{\mathbb{R}^2} f_{V\xi}\left(v, \xi\right) \mathbf{1}_{(\alpha-\mu^\tau, \beta-\mu^\tau)}\left(\sqrt{\mathbf{\Sigma}^\tau}\mathbf{U} + \mathbf{V}\right) \mathrm{d}v\, \mathrm{d}y}, \tag{A.3}$$

where $f_{V\xi}$ denotes the joint density function of V and ξ. When these two variables are independent, we we assume, the denominator is given by

$$\int_{\mathbb{R}^2} f_{V\xi}\left(v, \xi\right) \mathbf{1}_{(\alpha-\mu^\tau, \beta-\mu^\tau)}\left(\sqrt{\mathbf{\Sigma}^\tau}\mathbf{U} + \mathbf{V}\right) \mathrm{d}v\, \mathrm{d}y = F_Z\left(\beta\right) - F_Z\left(\alpha\right), \tag{A.4}$$

where Z is a normal distribution with mean μ^τ and variance $\Sigma^\tau + \Sigma^\xi$. It immediately follows that

$$F_Z\left(\beta\right) - F_Z\left(\alpha\right) = \Phi\left(\frac{\beta - \mu^\tau}{\sqrt{\Sigma^\tau + \Sigma^\xi}}\right) - \Phi\left(\frac{\alpha - \mu^\tau}{\sqrt{\Sigma^\tau + \Sigma^\xi}}\right) = \Phi\left(\beta'\right) - \Phi\left(\alpha'\right), \tag{A.5}$$

where $\alpha' \overset{\text{def}}{=} \frac{\alpha - \mu^\tau}{\sqrt{\Sigma^\tau + \Sigma^\xi}}$ and $\beta' \overset{\text{def}}{=} \frac{\beta - \mu^\tau}{\sqrt{\Sigma^\tau + \Sigma^\xi}}$ are linear transformations of the original interval boundaries. The numerator of the modified conditional expectation can be expressed as

$$\int_{\mathbb{R}^2} v\, f_{V\xi}\left(v, \xi\right) \mathbf{1}_{(\alpha-\mu^\tau, \beta-\mu^\tau)}\left(\sqrt{\mathbf{\Sigma}^\tau}\mathbf{U} + \mathbf{V}\right) \mathrm{d}v\, \mathrm{d}y = \int_{\mathbb{R}} v\, \phi\left(v\right) \int_{\alpha-\tau(v)}^{\beta-\tau(v)} \phi\left(\frac{\xi}{\sqrt{\Sigma^\xi}}\right) \mathrm{d}y\, \mathrm{d}v, \tag{A.6}$$

where, for notational convenience, we define $\tau\left(v\right) \overset{\text{def}}{=} \mu^\tau + \sqrt{\Sigma^\tau}v$. Evaluating the inner integral yields

$$\int_{\mathbb{R}} v\, \phi\left(v\right) \left(\Phi\left(\frac{\beta - \tau\left(v\right)}{\sqrt{\Sigma^\xi}}\right) - \Phi\left(\frac{\alpha - \tau\left(v\right)}{\sqrt{\Sigma^\xi}}\right)\right) \mathrm{d}v. \tag{A.7}$$

This expression is the difference of two integrals of the form

$$\int_{\mathbb{R}} v\,\phi\left(v\right)\Phi\left(c_0 - c_1 v\right) \tag{A.8}$$

for constants c_0 and $c_1 > 0$. Integration-by-parts yields

$$\int_{\mathbb{R}} v\,\phi\left(v\right)\Phi\left(c_0 - c_1 v\right) = -\frac{c_1 e^{-\frac{c_0^2}{2c_1^2+2}}}{\sqrt{2\pi}\sqrt{c_1^2+1}} = -\frac{c_1}{\sqrt{c_1^2+1}}\phi\left(\frac{c_0}{\sqrt{c_1^2+1}}\right). \tag{A.9}$$

Using this result, we can evaluate the numerator

$$\int_{\mathbb{R}^2} v\, f_{V\xi}\left(v,\xi\right)\mathbf{1}_{\left(\alpha-\mu^\tau,\beta-\mu^\tau\right)}\left(\sqrt{\mathbf{\Sigma}^\tau}\mathbf{U}+\mathbf{V}\right)\mathrm{d}v\,\mathrm{d}y = \sqrt{\frac{\Sigma^\tau}{\Sigma^\tau+\Sigma^\xi}}\left(\phi\left(\alpha'\right)-\phi\left(\beta'\right)\right). \tag{A.10}$$

Combining our expressions for the numerator and denominator yields the desired result

$$\mathbb{E}\left[\tau \mid \tau+\xi \in \left(\alpha,\beta\right)\right] = \mu^\tau + \sqrt{\Sigma^\tau}\cdot\sqrt{\frac{\Sigma^\tau}{\Sigma^\tau+\Sigma^\xi}}\cdot\frac{\phi\left(\alpha'\right)-\phi\left(\beta'\right)}{\Phi\left(\beta'\right)-\Phi\left(\alpha'\right)}. \tag{A.11}$$

A.2. Proof of Proposition 1

Starting from Equation 4.13, we have

$$\frac{\eta-1}{\eta}\alpha\frac{p_i y_i}{k_i} = e^{\tau_i} r \tag{A.12}$$

$$\frac{p_i y_i}{k_i} = e^{\tau_i} r \alpha^{-1}\frac{\eta}{\eta-1}. \tag{A.13}$$

Taking logs, we have

$$\log\frac{p_i y_i}{k_i} = \tau_i + \log r - \underbrace{\log\alpha}_{\frac{\partial \log y/k}{\partial k}} + \log\frac{\eta}{\eta-1}. \tag{A.14}$$

Taking conditional expectations, we have

$$\mathbb{E}\left[\log\frac{p_i y_i}{k_i} \mid D_i = j\right] = \log\frac{\eta}{\eta-1} + \log r - \frac{\partial \log y/k}{\partial k} + \mathbb{E}\left[\tau_i \mid T_{j-1} < \tau_i + \xi_i < T_j\right]. \tag{A.15}$$

Using Lemma 1 and Assumption 2, we have

$$\mathbb{E}\left[\tau_i \mid T_{j-1} < \tau_i + \xi_i < T_j\right] = \mu^\tau + \sqrt{\Sigma^\tau}\cdot\rho\cdot\lambda_j. \tag{A.16}$$

Using the definition of T_j

$$\Phi\left(\frac{T_j-\mu^\tau}{\sqrt{\Sigma^\tau}/\rho}\right) = \sum_{j'<j}\mathcal{P}_{j'} \tag{A.17}$$

$$T_j = \mu^\tau + \frac{\sqrt{\Sigma^\tau}}{\rho}\Phi^{-1}\left(\sum_{j'<j}\mathcal{P}_{j'}\right) \tag{A.18}$$

Using Assumption 4, we have

$$T_j = -\rho\sqrt{\Sigma^\tau}\lambda_0 + \frac{\sqrt{\Sigma^\tau}}{\rho}\Phi^{-1}\left(\sum_{j'<j}\mathcal{P}_{j'}\right) \tag{A.19}$$

where $\lambda_0 = -\phi\left(\Phi^{-1}\left(\mathcal{P}_0\right)\right)/\mathcal{P}_0$ is solely a function of the observed frequency of survey responses. It follows that

$$\Sigma^\tau = \left(\frac{T_j}{\frac{1}{\rho}\Phi^{-1}\left(\sum_{j'<j}\mathcal{P}_{j'}\right) - \rho\lambda_0}\right)^2 \tag{A.20}$$

Substituting this back into the conditional expectation yields

$$\mathbb{E}\left[\tau_i \mid T_{j-1} < \tau_i + \xi_i < T_j\right] = \mu^\tau + \frac{T_j}{\frac{1}{\rho^2}\Phi^{-1}\left(\sum_{j'<j}\mathcal{P}_{j'}\right) - \lambda_0}\cdot\lambda_j \tag{A.21}$$

$$= \mu^\tau + T_j\underbrace{\frac{\lambda_j}{\frac{1}{\rho^2}\Phi^{-1}\left(\sum_{j'<j}\mathcal{P}_{j'}\right) - \lambda_0}}_{\backslash mathcal\{X\}_j}. \tag{A.22}$$

In the special case of a binary indicator variable, as is the case in the EFIGE dataset, this can be further simplified, as there is only a single cutoff threshold T to consider

$$\mathbb{E}\left[\log\frac{p_i y_i}{k_i} \mid D_i = 0\right] = \log\frac{\eta}{\eta-1} + \log r - \frac{\partial\log y/k}{\partial k} + \mu^\tau + T\frac{\lambda_0}{\frac{1}{\rho^2}\Phi^{-1}\left(\mathcal{P}_0\right) - \lambda_0}, \tag{A.23}$$

A.3. Proof of Proposition 2

We start with the expression for aggregate output

$$Y = \left(\int y_i^{\frac{\eta-1}{\eta}}\,\mathrm{d}i\right)^{\frac{\eta}{\eta-1}} \tag{A.24}$$

Intermediate firms solve (z_i set to 1 $\forall i$)

$$\max_{k_i,\ell_i}\quad p_i y_i - e^{\tau_i} r k_i - w\ell_i, \qquad y_i = k_i^\alpha \ell_i^{1-\alpha} \tag{A.25}$$

Prices given by (P normalized to 1)

$$p_i y_i = y_i^{\frac{\eta-1}{\eta}} Y^{\frac{1}{\eta}} \tag{A.26}$$

The first order conditions of firm i are

$$Y^{\frac{1}{\eta}}\alpha\frac{\eta-1}{\eta}k_i^{\alpha\frac{\eta-1}{\eta}-1}\ell_i^{(1-\alpha)\frac{\eta-1}{\eta}} - e^{\tau_i} r = 0 \tag{A.27}$$

and

$$Y^{\frac{1}{\eta}}\left(1-\alpha\right)\frac{\eta-1}{\eta}k_i^{\alpha\frac{\eta-1}{\eta}}\ell_i^{(1-\alpha)\frac{\eta-1}{\eta}-1} - w = 0 \tag{A.28}$$

This implies an optimal factor input mix

$$\frac{\alpha}{1-\alpha}\frac{w}{re^{\tau_i}} = \frac{k_i}{\ell_i} \tag{A.29}$$

so the problem can be rewritten as

$$\max_{k_i} \quad Y^{\frac{1}{\sigma}} k_i^{\frac{\eta-1}{\eta}} \left(\frac{1-\alpha}{\alpha} \frac{re^{\tau_i}}{w} \right)^{(1-\alpha)\frac{\eta-1}{\eta}} - \frac{1}{\alpha} re^{\tau_i} k_i \tag{A.30}$$

Rewriting the first order conditions in terms of capital k_i yields

$$Y^{\frac{1}{\eta}} \frac{\eta-1}{\eta} k_i^{-\frac{1}{\eta}} \left(\frac{1-\alpha}{\alpha} \frac{re^{\tau_i}}{w} \right)^{(1-\alpha)\frac{\eta-1}{\eta}} = \frac{1}{\alpha} re^{\tau_i} \tag{A.31}$$

Thus, each intermediate firm i chooses

$$k_i = \left(\frac{\eta-1}{\eta} \right)^{\eta} \left(\frac{re^{\tau_i}}{\alpha} \right)^{-\alpha(\eta-1)-1} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)(\eta-1)} Y, \tag{A.32}$$

$$\ell_i = \left(\frac{\eta-1}{\eta} \right)^{\eta} \left(\frac{re^{\tau_i}}{\alpha} \right)^{-\alpha(\eta-1)} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)(\eta-1)-1} Y, \quad \text{and} \tag{A.33}$$

$$y_i = \left(\frac{\eta-1}{\eta} \right)^{\eta} Y \left(\frac{re^{\tau_i}}{\alpha} \right)^{-\alpha\eta} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)\eta} \tag{A.34}$$

By market clearing, aggregate capital is

$$\int_0^1 k_i \mathrm{d}i = \int \left(\frac{\eta-1}{\eta} \right)^{\eta} \left(\frac{re^{\tau_i}}{\alpha} \right)^{-\alpha(\eta-1)-1} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)(\eta-1)} Y \mathrm{d}i \tag{A.35}$$

$$K = \left(\frac{\eta-1}{\eta} \right)^{\eta} \left(\frac{r}{\alpha} \right)^{-\alpha(\eta-1)-1} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)(\eta-1)} Y \int_0^1 e^{\tau_i(-\alpha(\eta-1)-1)} \mathrm{d}i \tag{A.36}$$

Similarly, aggregate labor is

$$\int_0^1 \ell_i \mathrm{d}i = \int_0^1 \left(\frac{\eta-1}{\eta} \right)^{\eta} \left(\frac{re^{\tau_i}}{\alpha} \right)^{-\alpha(\eta-1)} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)(\eta-1)-1} Y \mathrm{d}i \tag{A.37}$$

$$1 = \left(\frac{\eta-1}{\eta} \right)^{\eta} \left(\frac{r}{\alpha} \right)^{-\alpha(\eta-1)} \left(\frac{w}{1-\alpha} \right)^{-(1-\alpha)(\eta-1)-1} Y \int_0^1 e^{\tau_i(-\alpha(\eta-1))} \mathrm{d}i \tag{A.38}$$

This implies

$$k_i = \frac{e^{\tau_i(-\alpha(\eta-1)-1)}}{\int_0^1 e^{\tau_i(-\alpha(\eta-1)-1)} \mathrm{d}i} K, \quad \text{and} \tag{A.39}$$

$$\ell_i = \frac{e^{\tau_i(-\alpha(\eta-1))}}{\int_0^1 e^{\tau_i(-\alpha(\eta-1))} \mathrm{d}i} \tag{A.40}$$

Substituting into the expression for aggregate output

$$Y^{\frac{\eta-1}{\eta}} = \int_0^1 \left(\left(\frac{e^{\tau_i(-\alpha(\eta-1)-1)}}{\int_0^1 e^{\tau_i(-\alpha(\eta-1)-1)} \mathrm{d}i} K \right)^{\alpha} \left(\frac{e^{\tau_i(-\alpha(\eta-1))}}{\int_0^1 e^{\tau_i(-\alpha(\eta-1))} \mathrm{d}i} \right)^{1-\alpha} \right)^{\frac{\eta-1}{\eta}} \mathrm{d}i \tag{A.41}$$

$$Y = K^{\alpha} \left(\frac{1}{\int_0^1 e^{\tau_i(-\alpha(\eta-1)-1)} \mathrm{d}i} \right)^{\alpha} \left(\frac{1}{\int_0^1 e^{\tau_i(-\alpha(\eta-1))} \mathrm{d}i} \right)^{1-\alpha} \left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))} \mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \tag{A.42}$$

Aggregate productivity (TFP) is therefore

$$\frac{Y}{K^{\alpha}}=\left(\frac{1}{\int_0^1 e^{\tau_i(-\alpha(\eta-1)-1)}\mathrm{d}i}\right)^{\alpha}\left(\frac{1}{\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i}\right)^{1-\alpha}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{\frac{\eta}{\eta-1}} \tag{A.43}$$

$$\mathrm{TFP}=\left(\frac{1}{\int_0^1 e^{\tau_i(-\alpha(\eta-1)-1)}\mathrm{d}i}\right)^{\alpha}\left(\frac{1}{\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i}\right)^{1-\alpha}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{\frac{\eta}{\eta-1}} \tag{A.44}$$

In logs

$$\log \mathrm{TFP}=-\alpha m_T(-\alpha\left(\eta-1\right)-1)-\left(1-\alpha\right)m_T(-\alpha\left(\eta-1\right))+\frac{\eta}{\eta-1}m_T(-\alpha\left(\eta-1\right)) \tag{A.45}$$

$$\log \mathrm{TFP}=-\alpha m_T(-\alpha\left(\eta-1\right)-1)+\left(\frac{\eta}{\eta-1}-\left(1-\alpha\right)\right)m_T(-\alpha\left(\eta-1\right)), \tag{A.46}$$

where $m_T\left(.\right)$ denotes the cumulant-generating-function of τ. Now we express first order conditions in terms of labor

$$Y^{\frac{\eta-1}{\eta}}=\int_0^1\left(\left(\frac{\alpha}{1-\alpha}\frac{w}{re^{\tau_i}}\right)^{\alpha}\ell_i\right)^{\frac{\eta-1}{\eta}}\mathrm{d}i \tag{A.47}$$

$$Y=\left(\frac{\alpha}{1-\alpha}\frac{w}{r}\right)^{\alpha}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{\frac{\eta}{\eta-1}}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{-1} \tag{A.48}$$

$$\log Y=\alpha\log\left(\frac{\alpha}{1-\alpha}\frac{w}{r}\right)+\frac{1}{\eta-1}m_T(-\alpha\left(\eta-1\right)) \tag{A.49}$$

Differencing against the first-best equation

$$\log Y-\log Y^{*}=\alpha\left(\log\omega-\log\omega^{*}\right)+\frac{1}{\eta-1}m_T(-\alpha\left(\eta-1\right)), \tag{A.50}$$

where $\omega\overset{\mathrm{def}}{=}\frac{w}{r}$ denotes the relative factor prices.

In equilibrium, aggregate output must be consistent with individual output choices. Starting from A.34, we have

$$y_i=\left(\frac{\eta-1}{\eta}\right)^{\eta}Y\left(\frac{re^{\tau_i}}{\alpha}\right)^{-\alpha\eta}\left(\frac{w}{1-\alpha}\right)^{-(1-\alpha)\eta} \tag{A.51}$$

Substituting into the expression for aggregate output

$$Y^{\frac{\eta-1}{\eta}}=\int_0^1 y_i^{\frac{\eta-1}{\eta}}\,\mathrm{d}i \tag{A.52}$$

gives

$$Y=Y\left(\frac{\eta-1}{\eta}\right)^{\eta}\left(\frac{1}{r}\right)^{\eta}\omega^{-(1-\alpha)\eta}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{\frac{\eta}{\eta-1}} \tag{A.53}$$

Substituting in Equation A.48:

$$r=\frac{\eta-1}{\eta}\alpha^{\alpha}\left(1-\alpha\right)^{1-\alpha}\omega^{-(1-\alpha)}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{\frac{1}{\eta-1}} \tag{A.54}$$

$$\omega^{1-\alpha} = \frac{\eta-1}{\eta}\alpha^{\alpha}(1-\alpha)^{1-\alpha} r^{-1}\left(\int_0^1 e^{\tau_i(-\alpha(\eta-1))}\mathrm{d}i\right)^{\frac{1}{\eta-1}} \tag{A.55}$$

In logs

$$\log\omega = \frac{1}{1-\alpha}\left(\frac{1}{\eta-1}m_T(-\alpha(\eta-1)) - \log r + \log\frac{\eta-1}{\eta} + \alpha\log\alpha + (1-\alpha)\log(1-\alpha)\right) \tag{A.56}$$

$$\log\omega - \log\omega^* = \frac{1}{1-\alpha}\frac{1}{\eta-1}m_T(-\alpha(\eta-1)) \tag{A.57}$$

Substitute into the equation for aggregate output

$$\log Y - \log Y^* = \frac{\alpha}{1-\alpha}\frac{1}{\eta-1}m_T(-\alpha(\eta-1)) + \frac{1}{\eta-1}m_T(-\alpha(\eta-1)) \tag{A.58}$$

$$\log Y - \log Y^* = \frac{1}{1-\alpha}\frac{1}{\eta-1}m_T(-\alpha(\eta-1)) \tag{A.59}$$

The representative firm solves

$$\max_K Y - RK, \tag{A.60}$$

where $Y = \text{TFP}\cdot K^{\alpha}$. Define aggregate MRPK as the MRPK of the representative firm, $\text{MRPK} \overset{\text{def}}{=} \frac{Y}{K}$.

$$\text{MRPK} = \frac{Y}{K} = \text{TFP}^{\frac{1}{\alpha}}Y^{1-\frac{1}{\alpha}}. \tag{A.61}$$

where we substitute the expression $(Y/\text{TFP})^{\frac{1}{\alpha}}$ for aggregate capital K. In logs,

$$\log\text{MRPK} = \frac{1}{\alpha}\log\text{TFP} + \left(1-\frac{1}{\alpha}\right)\log Y \tag{A.62}$$

In first-best

$$\log\text{MRPK}^* = \left(1-\frac{1}{\alpha}\right)\log Y^*. \tag{A.63}$$

Differencing against the first-best equation yields

$$\log\text{MRPK} - \log\text{MRPK}^* = \frac{1-\alpha}{\alpha}\log\frac{Y^*}{Y} + \frac{1}{\alpha}\log\text{TFP} \tag{A.64}$$

substitute in the expressions for $\log\frac{Y^*}{Y}$ and $\log\text{TFP}$:

$$\begin{aligned}\log\text{MRPK} - \log\text{MRPK}^* = &-\frac{1-\alpha}{\alpha}\frac{1}{1-\alpha}\frac{1}{\eta-1}m_T(-\alpha(\eta-1)) + \\ &+\frac{1}{\alpha}\left(-\alpha m_T(-\alpha(\eta-1)-1) + \left(\frac{\eta}{\eta-1}-(1-\alpha)\right)m_T(-\alpha(\eta-1))\right)\end{aligned} \tag{A.65}$$

simplifying

$$\begin{aligned}\log\text{MRPK} - \log\text{MRPK}^* = &-\frac{1}{\alpha}\frac{1}{\eta-1}m_T(-\alpha(\eta-1)) \\ &- m_T(-\alpha(\eta-1)-1) + \frac{1}{\alpha}\left(\frac{\eta}{\eta-1}-(1-\alpha)\right)m_T(-\alpha(\eta-1))\end{aligned} \tag{A.66}$$

$$\log \mathrm{MRPK} - \log \mathrm{MRPK}^* = m_T(-\alpha(\eta-1)) - m_T(-\alpha(\eta-1)-1) \tag{A.67}$$

In the case of a normally distributed τ, this becomes

$$\log \mathrm{MRPK} - \log \mathrm{MRPK}^* = -\alpha(\eta-1)\mu^\tau + \frac{1}{2}\alpha^2(\eta-1)^2\Sigma^\tau - \left((-\alpha(\eta-1)-1)\mu^\tau + \frac{1}{2}(\alpha(\eta-1)+1)^2\Sigma^\tau\right) \tag{A.68}$$

which simplifies to:

$$\log \mathrm{MRPK} - \log \mathrm{MRPK}^* = \mu^\tau - \frac{1}{2}(2\alpha(\eta-1)+1)\Sigma^\tau \tag{A.69}$$

Using the definition of aggregate MRPK,

$$\mathrm{MRPK} = \frac{Y}{K} = \mathrm{TFP} \cdot K^{\alpha-1} \tag{A.70}$$

we can write aggregate capital K as a function of TFP and aggregate MRPK

$$K = \left(\frac{\mathrm{MRPK}}{\mathrm{TFP}}\right)^{\frac{1}{\alpha-1}} \tag{A.71}$$

In logs,

$$\log \frac{K}{K^*} = \frac{1}{1-\alpha}\left(\log \frac{\mathrm{TFP}}{\mathrm{TFP}^*} - \log \frac{\mathrm{MRPK}}{\mathrm{MRPK}^*}\right). \tag{A.72}$$

Using the aggregate output equation

$$Y = \mathrm{TFP} \cdot K^\alpha, \tag{A.73}$$

we have

$$\log \frac{Y}{Y^*} = \log \frac{\mathrm{TFP}}{\mathrm{TFP}^*} + \alpha \log \frac{K}{K^*}. \tag{A.74}$$

Substituting,

$$\log \frac{Y}{Y^*} = \log \frac{\mathrm{TFP}}{\mathrm{TFP}^*} + \frac{\alpha}{1-\alpha}\left(\log \frac{\mathrm{TFP}}{\mathrm{TFP}^*} - \log \frac{\mathrm{MRPK}}{\mathrm{MRPK}^*}\right) \tag{A.75}$$

$$\log \frac{Y}{Y^*} = \frac{1}{1-\alpha}\log \frac{\mathrm{TFP}}{\mathrm{TFP}^*} - \frac{\alpha}{1-\alpha}\log \frac{\mathrm{MRPK}}{\mathrm{MRPK}^*} \tag{A.76}$$

For a normal distribution, this becomes

$$\log \frac{Y}{Y^*} = \frac{1}{1-\alpha}\frac{\alpha(\alpha\eta+1-\alpha)}{2} \cdot \Sigma^\tau - \frac{\alpha}{1-\alpha}\left(\mu^\tau + (2\alpha\eta+1-2\alpha)\,\Sigma^\tau\right). \tag{A.77}$$

As discussed in the main text, we offset the OAH effect term, thereby making $\log \frac{\mathrm{MRPK}}{\mathrm{MRPK}^*}$ depend solely upon μ^τ, and are left with our expression for $\Delta \log \mathrm{GDP}$,

$$\Delta \log \mathrm{GDP} = \frac{1}{1-\alpha}\frac{\alpha(\alpha\eta+1-\alpha)}{2} \cdot \Sigma^\tau - \frac{\alpha}{1-\alpha}\mu^\tau. \tag{A.78}$$

B. Additional Tables

Table 3: Estimates of the GDP cost of red tape, using alternative calibrations of ρ.

$\rho = 1/3$

Country	GDP Loss	Misallocation	Under-Investment
France	-6.30%	-3.46%	-2.95%
Italy	-0.75%	-0.26%	-0.49%
Spain	-0.28%	-0.10%	-0.18%
Germany	-0.14%	-0.06%	-0.08%
Hungary	-0.10%	-0.05%	-0.05%
Austria	-0.09%	-0.04%	-0.05%
United Kingdom	-0.08%	-0.04%	-0.04%

$\rho = 2/3$

Country	GDP Loss	Misallocation	Under-Investment
France	-2.90%	-0.56%	-2.36%
Italy	-0.89%	-0.15%	-0.74%
Spain	-0.38%	-0.07%	-0.31%
Germany	-0.20%	-0.05%	-0.15%
Hungary	-0.14%	-0.04%	-0.10%
Austria	-0.12%	-0.04%	-0.09%
United Kingdom	-0.11%	-0.03%	-0.08%

Table Notes: The table above presents country-level estimates of the cost of red tape analogous to those found in Table 2, where here the value of the parameter is calibrated to be $\frac{1}{3}$ in the upper panel, and $\frac{2}{3}$ in the lower one.

Table 4: Unadjusted Estimates of the GDP cost of red tape

Country	% Constrained	GDP Loss	Misallocation	Under-Investment
France	45.8%	-1.96%	-1.29%	-0.69%
Italy	26.2%	-0.52%	-0.19%	-0.33%
Spain	14.2%	-0.20%	-0.08%	-0.12%
Germany	7.8%	-0.09%	-0.05%	-0.04%
Hungary	5.4%	-0.05%	-0.04%	-0.01%
Austria	4.8%	-0.05%	-0.04%	-0.01%
United Kingdom	4.3%	-0.04%	-0.04%	-0.00%

Table Notes: The table above presents country-level estimates of the cost of red tape. "% Constrained" is the proportion of that country's firms which report being constrained by bureaucratic regulation. "GDP Loss" denotes the total GDP loss, while "Misallocation" and "Under-Investment" decompose the total loss into the components stemming from misallocation and underinvestment, respectively. "GDP Loss" is calculated using Equation A.77. "Misallocation" is calculated using Equation 4.17. "Under-Investment" is calculated using Equation 4.18 and does not include the impact of a flat rate tax to offset the Oi-Hartman-Abel effect. Note that GDP Loss is not exactly the sum of the Misallocation and Under-Investment columns because the effects are additive in logs, not levels.

Table 5: GDP cost of red tape, using T estimated on a subset of the EFIGE dataset.

Country	% Constrained	GDP Loss	Misallocation	Under-Investment
France	45.8%	-3.73%	-1.18%	-2.58%
Italy	26.2%	-0.78%	-0.18%	-0.61%
Spain	14.2%	-0.31%	-0.08%	-0.24%
Germany	7.8%	-0.16%	-0.05%	-0.11%
Hungary	5.4%	-0.11%	-0.04%	-0.07%
Austria	4.8%	-0.10%	-0.03%	-0.06%
United Kingdom	4.3%	-0.19%	-0.03%	-0.06%

Table Notes: The table above presents country-level estimates of the cost of red tape analogous to those found in Table 2, where here the threshold T is estimated using a subset of the EFIGE dataset. Here, we use only firm-level financials from France, Hungary, Italy, and Spain. These countries have high quality coverage in the Amadeus dataset, and thus are less susceptible to problems of sample non-representativeness.

Table 6: MRPK and Bureaucratic Constraints

	log MRPK			
	(1)	(2)	(3)	(4)
Bureaucracy Constrained	0.062**		0.054*	
	(2.19)		(1.91)	
$\mathcal{X}$		0.086***		0.072***
		(3.23)		(2.73)
Financial Constraint F.E	No	No	Yes	Yes
Firm Size F.E.	No	No	Yes	Yes
Country F.E.	Yes	Yes	Yes	Yes
Sector F.E.	Yes	Yes	Yes	Yes

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table Notes: The table above presents regression estimates of the log of MRPK on firms' survey responses (in column 1) and firms' $\mathcal{X}$. $\mathcal{X}$ is defined as in Proposition 1. In our table, we omit the regression constant because it is not separately identifiable from the country and sector fixed effects. Financial Constraint fixed effects control for whether the firms reported that Financial Constraints were a major constraint on growth. Firm Size fixed effects control for the number of employees at the firm. t-statistics based on Huber-White standard errors are reported in parentheses.

Table 7: MRPL and Bureaucratic Constraints

	log MRPL	
	(1)	(2)
Bureaucracy Constrained	-0.005	
	(-0.34)	
$\mathcal{X}$		-0.010
		(-0.66)
Country F.E.	Yes	Yes
Sector F.E	Yes	Yes

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table Notes: The table above presents regression estimates of the log of MRPL on firms' survey responses (in column 1) and $\mathcal{X}$, which is defined as in Proposition 1. In our table, we omit the regression constant because it is not separately identifiable from the country and sector fixed effects. t-statistics based on Huber-White standard errors are reported in parentheses.