Tax Heterogeneity and Misallocation*

Barış Kaymak† Immo Schott‡

December 4, 2025

Abstract

There is substantial asymmetry in effective corporate income tax rates across firms. While tax asymmetries reduce productivity in frictionless economies, they can improve efficiency in distorted economies by alleviating other frictions. We develop a framework to estimate the productivity effects of tax asymmetries in distorted economies. Using U.S. firm-level balance sheet data alongside measures of marginal tax rates, we find a positive correlation between tax rates and factor productivity, suggesting that tax asymmetries *exacerbate* other economic frictions. Our model implies significant aggregate productivity gains from eliminating tax asymmetries.

*The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

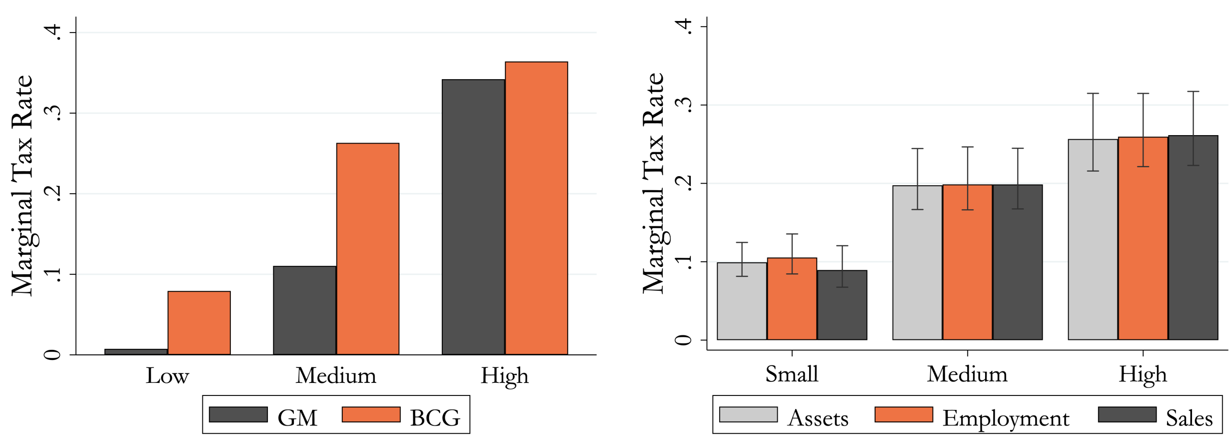
†Department of Economic Research, Federal Reserve Bank of Cleveland, 1455 E 6th St, Cleveland, OH 44114, United States. E-mail: barkaymak@gmail.com

‡Federal Reserve Board. Division of International Finance, 20th and C Streets, NW, Washington, DC 20551, United States. E-mail: immoschott@gmail.com

1 Introduction

In efficient economies, marginal products of inputs are equalized across firms. Yet, numerous studies following Hsieh and Klenow (2009) document substantial productivity differences across firms, suggesting significant gains from resource reallocation. The sources of input misallocation, however, remain elusive, preventing policy guidance.

We examine whether corporate taxation contributes to misallocation. Although the tax code does not distinguish individual firms *de jure*, provisions for deductions and allowances, such as imperfect loss offset or interest deduction, create *de facto* asymmetries in effective marginal tax rates (EMTR) across firms. This can cause factor productivity to diverge between firms and reduce efficiency. In our sample of U.S. firms, EMTR varies widely across firms, with large firms typically facing higher rates (Figure 1). Because these firms represent the majority of economic activity, the efficiency implications of tax asymmetries are potentially large, especially if size indicates productivity.



(a) Tax rates by tax tercile

(b) Tax rates by size class

Figure 1: Asymmetry in Effective Marginal Corporate Tax Rates

Note.- GM and BCG refer to the rates estimated by Graham and Mills (2008) and Blouin, Core, and Guay (2010). Figure shows average marginal corporate tax rates by terciles of tax rates (panel a), and by terciles of firm size as measured by assets, employment, or sales (panel b). The high and low values shown with whiskers correspond respectively to the two tax rate estimates by GM and BCG. Bars show their average. Source: Compustat. Sample period is 1980–2016.

Whether tax asymmetries reduce efficiency is theoretically ambiguous. Although they impair efficiency in competitive, otherwise frictionless economies, they might actually *improve* efficiency in distorted economies by offsetting existing frictions. For instance, interest deductions might ease credit constraints and raise investment, while loss-offset provisions might prevent premature bankruptcy among liquidity-constrained firms by reducing their tax obligations at times of distress. Whether this is the case is an empirical question.

In pursuit of an answer, we study the effect of tax asymmetries on total factor productivity (TFP) in a production model with heterogeneous firms and distortions to factor allocation. We derive analytical formulas that relate the joint cross-sectional distribution of EMTRs and factor productivities to aggregate TFP. The key insight is that the relationship between corporate taxation and other distortions is captured by the correlation between EMTR and productivity across firms. A positive correlation implies that high-productivity firms are taxed more heavily, lowering TFP.

Using firm-level data on effective marginal tax rates, we estimate that eliminating tax asymmetries would *raise* TFP by around 4%, holding technology and other distortions fixed. These gains arise from the empirically observed positive correlation between EMTRs and firm productivity: tax homogenization lowers EMTRs for more productive firms and reallocates capital toward them, thereby raising TFP. Because capital and labor are complementary, this capital reallocation also induces labor reallocation, which further increases TFP given the strong positive correlation between EMTRs and labor productivity. If we abstract from these correlations, the implied productivity gains from tax homogenization are below 1%, with the exception of the pre-1986 period. This indicates that tax asymmetries amplify rather than offset existing misallocation.

We then discuss forces that can modify both the magnitude and the interpretation of the productivity gains from tax homogenization. First, we trace out the dynamic implications of these productivity gains and show that induced capital accumulation and revenue-neutral tax adjustments further amplify long-run output beyond the static TFP improvements. Second, we allow for interactions between tax distortions and other frictions—such as imperfect deductibility of labor costs and financial constraints—and show that these interactions can substantially amplify the productivity gains when all observed correlations are interpreted causally. Third, the results are robust to major macroeconomic developments over our sample period, including the secular decline in the labor share of income. Finally, we find that the TFP improvements arise predominantly from more efficient resource allocation within sectors, indicating that tax heterogeneity mainly distorts competition among firms in the same industry.

This paper connects to the literature on the macroeconomic effects of factor misallocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). Papers in this literature typically adopt one of two approaches. The first approach measures the dispersion in marginal products and then attributes this dispersion to firm-specific distortions. Comparing the distorted economy to a model-implied frictionless environment, papers following this approach typically find large potential gains from removing frictions (see e.g., Hsieh and Klenow (2009), Gopinath, Kalemli-Özcan, Karabarbounis, and Villegas-Sanchez (2017), and Adamopoulos, Brandt, Leight, and Restuccia (2022)). A challenge in this literature has been to distinguish differences in marginal products from differences in production technologies or from measurement error (Bils, Klenow, and Ruane, 2021).[1]

The second approach measures frictions at the firm-level and computes the associated reductions in aggregate productivity (Midrigan and Xu, 2014; David and Venkateswaran, 2019; Kim, 2023; Stefanski, Ahlvik, Andersen, Harding, and Trew, 2025). Because of computational and theoretical complexities, this approach is limited in its ability to model and compute a large number of distortions and their interactions. Because specific frictions only partly explain the observed heterogeneity in marginal products across firms, most of these studies find limited productivity gains from reallocation.

We bridge these approaches by focusing on a specific friction – corporate tax asymmetries – while explicitly allowing for other, more general, distortions. In the spirit of Lipsey and Lancaster's (1956) analysis of the 'second best', removing tax heterogeneity in a frictional environment need not increase aggregate TFP, in contrast to most previous work, where eliminating distortions improves productivity by design. In that sense, our approach is closer to Harberger (1964), and, more

[1]Hsieh and Klenow (2009) address these issues by benchmarking their misallocation measure to the U.S. when assessing the extent of resource misallocation in India. Bils et al. (2021) use panel data to purge the firm-level dispersion in productivity of measurement error. Because our findings rely on the covariance between tax wedges and productivity, errors in specification or measurement of marginal products are less of a concern, see Section 3.

recently, Baqaee and Farhi (2020), Qian and Vereshchagina (2024), and Klenow, Pastén, and Ruane (2025), who respectively study productivity effects of price markups, financial constraints, and carbon taxes in distorted economies.[2] More broadly, our methodology provides a general framework for quantifying the allocative efficiency costs of any measurable distortion, extending beyond corporate taxation to other policies or frictions that create wedges in firms' factor costs.

Our paper also relates to the literature measuring the effects of corporate taxes on economic outcomes (Harberger, 1962; Hall and Jorgenson, 1967). This literature has studied specific provisions in the tax code, such as loss-offset provisions (Auerbach, 1986; Kaymak and Schott, 2019), investment tax credits (Cummins, Hassett, and Hubbard, 1994), jurisdictional differences (Djankov, Ganser, McLiesh, Ramalho, and Shleifer, 2010; Slattery and Zidar, 2020), as well as tax avoidance (Glover and Levine, 2023). Whereas these papers focus on the economic effects of the *level* of the corporate tax rate, we study the *dispersion* in tax rates and its implications for aggregate productivity.

2 Model

Consider a firm i that uses capital and labor to produce output with a decreasing-returns-to-scale production technology, $y^i = z^i F(k^i, n^i)$. Markets are competitive.[3] A firm's pre-tax profits are given by $\Pi^i_{\text{pre-tax}} = z^i F(k^i, n^i) - \omega^i_N n^i - \omega^i_R k^i$, where ω^i_N is a firm's marginal cost of labor and $\omega^i_R = r^i + \delta^i$ is the marginal costs of capital, consisting of the real rental rate r^i and the depreciation rate δ^i.[4]

Profits are taxed at firm-specific rates. Although the tax code does not distinguish individual firms *de jure*, effective marginal tax rates can differ across firms. This reflects special provisions

[2] In Klenow et al. (2025), a unilateral carbon tax can generate a differential effect across firms due to differences in fuel use intensity. Implications for aggregate productivity depend on the correlation between fossil fuel use and revenue productivity, which reflects other output distortions in the economy.

[3] This is equivalent to an economy with constant-returns-to-scale production and monopolistic competition as in Hsieh and Klenow (2009).

[4] In Appendix A, we demonstrate how capital adjustment costs and credit constraints can generate firm-specific marginal costs.

that shape the tax base, in particular loss provisions and various deductions and allowances. Loss offsets—which allow firms to carry losses forward or back to offset taxable income in other periods—tie current taxable income to a firm's past performance and give rise to firm-specific income tax rates τ_b^i.

When production expenses are not fully deductible from income, as is the case for capital costs, corporate income taxes can distort factor demand. The effective distortions to capital and labor are determined by the income tax rate τ_b^i and any deductions from taxable income; we illustrate this using depreciation allowances for capital and full expensing of labor costs.[5] Depreciation allowances determine the timing of cost recovery for tax purposes. Let d_{it} denote firm i's tax-depreciation allowance in year t, defined as the fraction of the asset's initial acquisition cost that the tax code permits the firm to deduct from taxable income in that year.[6] Following Hall and Jorgenson (1967), denote the present value of these depreciation allowances as

$$\lambda^i = \sum_{t=0}^{\infty} \frac{d_{it}}{(1+r_i)^t}. \tag{1}$$

A higher λ^i corresponds to more front-loaded depreciation for tax purposes and therefore to a larger tax shield per unit of capital.

The tax base is then given by $z^i F(k^i, n^i) - \omega_N^u n^i - \omega_R^i \lambda^i k^i$. The firm's problem consists of choosing k and n to maximize after-tax profits.

$$\max_{k,n} (1-\tau_b^i)\left[z^i F(k,n) - \omega_N^i n\right] - \omega_R^i \left(1 - \tau_b^i \lambda^i\right) k \tag{2}$$

The optimality condition for labor is

$$z^i F_n(k,n) = \omega_N^i,$$

[5] In Section 5, we consider imperfect expensing of labor costs. Appendix A shows a more comprehensive case that also includes debt financing.

[6] Tax-depreciation schedules generally differ from economic depreciation δ. The particular formulation of d_{it} depends on which methods were allowed that year (e.g., straight-line vs. double declining balance), the tax lives of capital types (e.g., structures vs. equipment), or the accounting method used (e.g. whether inventories of the firm follow last-in-first-out or first-in-first-out accounting). See Hall and Jorgenson (1967) and Gravelle (2001) for details.

which equates marginal product of labor to its marginal cost. Because labor costs are fully deductible, corporate taxation distorts employment only indirectly, through the distortion to capital. The optimality condition for capital is

$$z^i F_k(k, n) = \omega_R^i \cdot \underbrace{\frac{1 - \lambda^i \tau_b^i}{1 - \tau_b^i}}_{\omega_\tau^i} \equiv \omega_K^i, \tag{3}$$

where $z^i F_k(k, n)$ is the marginal product of capital and ω_K^i denotes the tax-adjusted marginal cost of capital. The term ω_K^i is composed of a pre-tax component ω_R^i and a tax distortion ω_τ^i, which embeds all the information about the effective rate before deductions τ_b^i and the depreciation allowance λ^i. The condition in (3) can equivalently be derived by formulating the firm's problem in (2) using an *effective* corporate income tax rate $\tau^i = 1 - 1/\omega_\tau^i$ that consolidates all deductions—such as loss offsets and depreciation allowances—into a single wedge.[7] Estimates of that effective tax rate were shown in Figure 1.

Equation (3) shows how distortions and tax asymmetries might affect aggregate productivity. Differences in ω_N^i or ω_K^i across firms are known to reduce aggregate productivity by inducing misallocation of inputs. We focus here on the allocative role of tax asymmetries, i.e., differences in ω_τ^i. That role generally depends on how ω_τ^i interacts with other distortions. For instance, if firms had common values of ω_N^i and ω_R^i, then tax asymmetries would distort the capital allocation and reduce TFP. But, if ω_R^i differed across firms and tax rates were $\omega_\tau^i = 1/\omega_R^i$ instead, taxes would offset capital distortions and improve TFP.

2.1 Tax heterogeneity and aggregate productivity

To focus on the marginal effect of tax heterogeneity on TFP, we consider a hypothetical economy where all firms face the same marginal tax rate. Our aim is to compare the output in that hypothetical economy using the same aggregate quantities of capital and labor as in the benchmark

[7]The effective tax rate in this case is given by $\tau^i = 1 - (1 - \tau_b^i)/(1 - \tau_b^i \lambda^i)$. In Appendix A, we also show the effective tax rate when capital is partly financed with debt.

economy. We leave the distribution of other distortions intact. In that sense, we adopt an agnostic view of the sources of differences in ω_R^i and ω_N^i. In Section 5, we consider the possibility that tax policy interacts with the underlying distortions and propose bounds on how those interactions may affect our analysis of aggregate productivity.

We begin by formulating TFP in the distorted economy. We now drop the firm subscript i for readability. We set $F(k,n) = zk^\alpha n^\beta$ with $\alpha + \beta = \gamma < 1$, and denote the joint distribution of firm-level TFP, capital and labor distortions by $G(z, \omega_K, \omega_N)$. The following proposition gives productivity in the distorted equilibrium.[8]

Proposition 1. *TFP in the distorted economy is*

$$Z = \frac{Y}{K^\alpha N^\beta} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{\beta}{1-\gamma}} \omega_K^{-\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{\beta}{1-\gamma}} \omega_K^{-\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{1-\alpha}{1-\gamma}} \omega_K^{-\frac{\alpha}{1-\gamma}} dG\right]^\beta}. \tag{4}$$

Aggregate TFP depends on the joint distribution of firm productivity z, and distortions ω_N and ω_K. The key is the relative distortion across firms rather than the level of distortions: a proportional change in distortions does not affect TFP.

In the hypothetical economy, all firms continue to face the same distortions to employment, ω_N, but capital distortions are altered to reflect a common tax wedge: $\omega_K' = \omega_R \cdot \bar{\omega}_\tau = \omega_K \cdot \bar{\omega}_\tau / \omega_\tau$ for some level of $\bar{\omega}_\tau$. The counterfactual TFP without tax heterogeneity is obtained by replacing ω_K by $\omega_{K'}$ in equation (4), giving the proposition below.

Proposition 2. *TFP in the counterfactual economy with homogeneous tax rates across firms is*

$$Z^* = \frac{\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{\beta}{1-\gamma}} \omega_K^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{\beta}{1-\gamma}} \omega_K^{-\frac{1-\beta}{1-\gamma}} \omega_\tau^{\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{1-\alpha}{1-\gamma}} \omega_K^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\alpha}{1-\gamma}} dG\right]^\beta}. \tag{5}$$

Because proportional changes in distortions do not affect aggregate TFP, $\bar{\omega}_\tau$ is absent in equation (5), and, therefore, the hypothetical economy is equivalent to an economy where tax wedge is removed altogether: $\omega_k' = \omega_R$.

[8] Proofs and derivations are in Appendix A.

The ratio Z^*/Z gives the marginal effect of eliminating tax heterogeneity on aggregate productivity. From equation (5), $Z^* = Z$ when ω_τ is constant across firms and $Z^* > Z$ whenever ω_τ is heterogeneous but orthogonal to other distortions.[9] More generally, however, if tax rates *correlate* with other distortions across firms, Z^* can be higher or lower than Z.

To gain further insights, consider an economy where distortions are distributed jointly according to a log-normal density. The resulting formulas help form an intuition about the sources of misallocation from tax heterogeneity and show how interactions between tax rates and other distortions can play an important role in assessing the allocative effects of tax distortions. Under the log-normality assumption, TFP in the distorted economy is given by

$$\ln Z = \ln \int z^{\frac{1}{1-\gamma}} dG - \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)\sigma_k^2 + \beta(1-\alpha)\sigma_n^2 + 2\alpha\beta\sigma_{kn}\right], \tag{6}$$

where σ_k^2 and σ_n^2 denote the variances of $\ln \omega_K$ and $\ln \omega_N$ respectively, and σ_{kn} is their covariance. Aggregate TFP reflects the underlying distribution of micro-level productivity levels, z, adjusted for efficiency losses caused by input distortions. We are interested in the change in TFP from eliminating the heterogeneity in tax rates. This implies setting $\sigma_k^2 = \sigma_R^2$ and $\sigma_{kn} = \sigma_{Rn}$ in equation (6). Taking differences and rearranging terms gives the next proposition.

Proposition 3. *If ω_K, ω_N, and ω_τ are jointly log-normally distributed, then eliminating tax heterogeneity yields the following change in aggregate TFP:*

$$\ln \frac{Z^*}{Z} = \frac{1}{2}\frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2 + \frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1) + \frac{\alpha\beta}{1-\gamma}\sigma_\tau^2 L_{n\tau}, \tag{7}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_K$ and $\ln \omega_N$ on $\ln \omega_\tau$.

Equation (7) has three distinct components. The first one is a pure misallocation component, representing the reduction in aggregate output caused by a dispersion in tax rates across firms.

[9] This follows from Jensen's inequality combined with the fact that $1-\beta > \alpha$.

In the absence of other economic distortions, or if such distortions were orthogonal to tax rates, this would be the improvement in TFP that can be expected from equalizing marginal tax rates. Productivity gains are increasing in the variance of tax wedges, σ_τ^2, the span of control parameter γ, and the capital share α.[10]

The second term captures the correlation of tax distortions with other capital distortions. If tax differences alleviate other capital distortions, then $\ln \omega_R$ and $\ln \omega_\tau$ are negatively correlated, and $L_{k\tau} < 1$. Eliminating tax differentials in this case need not generate TFP gains. Of course, taxes could also exacerbate the existing distortions. If firms with higher capital costs also face higher tax rates, then $L_{k\tau} > 1$, which raises the gains from eliminating tax differentials.

The last term captures the correlation of tax distortions with labor distortions. Equalization of tax rates lowers the tax rates for some firms and raises their capital demand. Because capital and labor are complements in production, labor is also reallocated toward those firms. This improves efficiency only if the marginal product of labor, ω_N, is higher at those firms on average, i.e., if $L_{n\tau} > 0$.

In our quantitative analysis, we implement equations (4) and (5) to estimate the productivity gains and use equation (7) to understand the sources of those gains.

3 Data and methodology

Our main data source is the annual Compustat database from 1980–2021, which provides balance-sheet data on publicly listed companies (Standard & Poor's, 2023).[11] Our formulas require information on employment, capital and output. For each firm, we define output as the sum of sales and changes in inventories during the year. We construct a firm's capital stock from its investment expenditures using the perpetual inventory method.

[10] To see this, note that $\frac{\alpha(1-\beta)}{1-\gamma} = \alpha + \frac{\alpha^2}{1-\gamma}$.

[11] See Appendix B for details.

3.1 Measuring tax distortions

We merge the Compustat data with estimates of corporate income tax rates for each firm-year from Graham and Mills (2008) and from Blouin et al. (2010). These studies simulate expected effective marginal tax rates at the firm-year level, τ_{it}, based on the corporate income tax code in effect during that year. They allow for provisions such as investment tax credits, the alternative minimum tax, capital expensing rules, and tax carryovers. This information is then combined with company-level balance sheet data from Compustat, e.g., detailed capital expenditures and past tax carryovers. Each study then estimates a stochastic process for a company's income to predict the applicable tax rates on future income associated with that year's investment. They then report the average marginal tax rate across possible realizations of income (see Graham (1996a) and Graham (1996b) for details). The simulated rates reflect the expected marginal tax rate on investment for each firm in a given year. Graham and Mills (2008) show that the simulated taxes closely approximate the actual taxes from tax records.

The simulated tax rates are generally below the statutory rates and vary significantly across firms. Appendix Table A1 summarizes the two EMTR measures. The average values for the two measures, denoted τ^{GM} and τ^{BCG}, are 17 and 24 percent whereas the statutory rate exceeds 34 percent for most firms.[12] The standard deviations of τ^{GM} and τ^{BCG} are 17 and 14 percent, mostly attributable to cross-sectional differences. While the two simulated rates are different, they are highly correlated ($\rho = 0.61$). We deploy both measures in our analysis below.

Because the two distinct tax measures are correlated imperfectly, we treat each as a potentially erroneous measure of the true EMTR. This allows us to leverage their empirical content by focusing on their common component. Specifically, we interpret the tax distortion associated with each measure as:

$$\ln \omega_{i\tau}^{*} = \ln \omega_{\tau} + \epsilon_i,$$

where ϵ_i is a classical measurement error with variance $\sigma_{\epsilon,i}^2$ and conditional mean $\mathbb{E}[\epsilon_i|\omega_\tau] = 0$ for

[12]For firms with more than $75,000 in income, the federal statutory rate varies from 34 to 51 percent during our sample period depending on income bracket and tax year.

$i \in \{$GM, BCG$\}$.

Measurement error in tax wedges exaggerates tax asymmetries and attenuates the estimated relationship between tax wedges and capital and labor productivity. These biases lead to underestimation of the TFP gain, as the next proposition states, and distort the assessment of its components in equation (7).

Proposition 4. *Assume that the tax wedge is measured with error, $\omega_\tau^* = \omega_\tau + \epsilon$ with $\mathbb{E}(\epsilon|\omega_\tau) = 0$. Then, replacing ω_τ by ω_τ^* in equation (7) underestimates the net TFP gain from eliminating tax heterogeneity:*

$$\ln(Z^*/Z)|\omega_\tau^* = \ln(Z^*/Z)|\omega_\tau - \frac{\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}. \tag{8}$$

To address measurement error, we treat ϵ_1 and ϵ_2 as independent. We find this reasonable because the two measures differ mainly in their assumptions regarding the underlying income process: whereas Graham and Mills (2008) assume a random walk for future returns, Blouin et al. (2010) allow for mean reversion. These assumptions affect the EMTRs via their predictions for the expected income volatility conditional on past performance. For instance, a firm with offsets for past losses would expect to pay taxes on investment only if future income is high enough to exceed its offset allowance, which is more likely if the predicted volatility is high. Thus, deviations from the true EMTR in a given year can be interpreted as errors in how each measure forecasts income volatility, relative to the volatility perceived by the firm for tax purposes. Our assumption interprets those prediction errors as orthogonal to each other. Another potential concern for measurement could be reporting errors in Compustat's balance sheet data. Because both studies use the same data source, this would violate our orthogonality assumption as they would bias both measures in the same direction, generating a positive correlation between ϵ_1 and ϵ_2. To the extent that ϵ_1 and ϵ_2 stem from reporting errors, however, our results below should be interpreted as conservative estimates of aggregate TFP gains given Proposition 4.

Formally, when we use equation (7), we estimate the variance of tax wedges with the covariance of the two measures: $\hat{\sigma}_\tau^2 = cov(\ln \omega_{1\tau}^*, \ln \omega_{2\tau}^*)$. To estimate $L_{n\tau}$ and $L_{k\tau}$, we regress labor and capital productivity on the measured tax wedges by instrumenting one measure with the other. This yields consistent estimates when the errors are uncorrelated across measures.

To accommodate potential deviations of tax wedges from log-normality, we also use the general formulas in equations (4) and (5) to compute the potential productivity gains. In particular, we partition the firms into quantile bins based on their marginal tax rate in each year. For each bin-year cell, we then calculate the average tax wedge and average firm-level TFP, along with capital and labor productivity as described above. In the Appendix we show that the resulting bias from measurement error aligns with the error in equation (8) when the distribution of marginal products is log-normal, conditional on the tax rate, regardless of the marginal distribution of tax rates. Accordingly, we adjust the estimated TFP gain for measurement error using equation (8) for each year. This requires the variance of the measurement error, which we estimate by $\hat{\sigma}^2_{\epsilon,jt} = var(\ln \omega^*_{j\tau}) - \hat{\sigma}^2_\tau$ for $j \in \{1, 2\}$.

3.2 Correlations between productivity and tax distortions

To measure capital and labor distortions, we use the optimality conditions for factor demand, $\ln \omega_K = \ln \alpha + \ln(y/k)$ and $\ln \omega_N = \ln \beta + \ln(y/n)$, where factor shares α and β are assumed common to all firms in an industry during a given year. Specifically, we estimate

$$\ln(y/k)_{it} = D^k_{st} + L_{k\tau,t} \ln \omega^*_{\tau,it} + e^k_{it}, \tag{9}$$

where i denotes the firm, and t the year. D^k_{st} are indicators for a full set of sector and year interactions. These indicators capture variations in capital shares and average distortions across sectors and years. Therefore, $L_{k\tau,t}$ reflects the correlation between the tax wedge and other capital distortions across firms in a given year.

Similarly, we estimate the cross-sectional correlation between labor distortions and tax wedges by:

$$\ln(y/n)_{it} = D^n_{st} + L_{n\tau,t} \ln \omega^*_{\tau,it} + e^n_{it}, \tag{10}$$

Because the OLS estimates of $L_{k\tau}$ and $L_{n\tau}$ are potentially attenuated by measurement error, we estimate equations (9) and (10) with an IV approach, where one tax measure is used as an instrument for the other.

To summarize the correlation patterns between different distortions, we report a pooled estimate for all years in Table 1. Columns 1-3 show the OLS estimates of (9) and (10). Panel A uses tax measures from Graham and Mills (2008), and Panel B uses the tax measures from Blouin et al. (2010). Both measures suggest a strong positive relationship between taxes and the three productivity measures, although they somewhat disagree on the strength of that relationship.

The implied correlation between tax wedges and other capital distortions (Column 1) differs across the two measures. Recall from (7) that a value below (above) one indicates that the tax wedge is negatively (positively) correlated with other capital distortions. While Graham and Mills's EMTR estimates suggest tax wedges and capital distortions correlate negatively, Blouin et al.'s estimates suggest they are orthogonal. By contrast, because both coefficients are larger than zero, the two measures suggest a positive correlation between labor distortions and the tax wedge (Column 2). This implies that reallocating labor toward high-tax firms would raise productivity. Column 3 shows the projections of firm-level TFP on the tax wedge, which we use below for some of our results. Consistent with the patterns from average labor and capital productivity, these estimates show that productive firms on average face higher marginal tax rates.

Correcting for the attenuation bias depicts a more consistent picture across measures (Columns 4-6). Both tax measures now indicate that tax wedges correlate positively with other capital distortions, implying that the heterogeneity in tax rates exacerbates the existing capital distortions.

With two tax measures, we have two estimates of the same underlying parameter for each year. We combine them by taking an average, weighted by the inverse of their variance. For example, using the estimates for the entire sample period, reported in columns 4 and 5 of Table 1, we obtain a value of 1.41 (s.e. 0.04) for $L_{k\tau}$ and 1.38 (s.e. 0.02) for $L_{n\tau}$ for the entire sample period.[13]

[13] Specifically, the variance minimizing weights are $\lambda_1 = var(b^{IV}_{k\tau,2})/(var(b^{IV}_{k\tau,1}) + var(b^{IV}_{k\tau,2}))$ for $b^{IV}_{k\tau,1}$, and $\lambda_2 = 1 - \lambda_1$ for $b^{IV}_{k\tau,2}$. For $L_{k\tau}$, for instance, the values in column (4) of Table 1 imply a weight of $0.44 = 0.08^2/(0.08^2 + 0.09^2)$ for 1.59 and 0.56 for 1.26, giving a weighted average of 1.41. The standard error for the weighted estimate is given by $var(b^{IV}_{k\tau,1}) \times var(b^{IV}_{k\tau,2})/(var(b^{IV}_{k\tau,1}) + var(b^{IV}_{k\tau,2}))$.

Table 1: Tax distortions and factor productivity

	(1) $\ln y/k$	(2) $\ln y/n$	(3) $\ln z$	(4) $\ln y/k$	(5) $\ln y/n$	(6) $\ln z$
			Panel A			
$\ln \omega_{1\tau}$	0.67	0.55	0.56	1.26	1.61	1.41
	(0.04)	(0.03)	(0.03)	(0.08)	(0.06)	(0.05)
			Panel B			
$\ln \omega_{2\tau}$	0.96	1.23	1.07	1.59	1.31	1.33
	(0.06)	(0.04)	(0.04)	(0.09)	(0.06)	(0.06)

Note.– The table shows the results from regressions of firm-level capital, labor, and total factor productivity on the tax wedge $(1/(1-\tau))$. Columns 1 to 3 report OLS estimates. Columns 4 to 6 report instrumental variable estimates to correct for measurement error. All specifications control for a full interaction of sector and year indicators. Standard errors are clustered at the firm level. The first-stage F-statistic for the IV regressions in Columns 4-6 is 49,592 (p=0.00). Data on productivity come from authors' calculations from Compustat. Data on marginal tax rates come from Graham and Mills (2008) in Panel A and Blouin et al. (2010) in Panel B.

Before we turn to the implications of our estimates for aggregate TFP and output, we need to set values for the parameters of the production function. For our baseline results, we keep those parameters fixed over time. This allows us to highlight the changes in the interactions between distortions when presenting the time trends. We set $\gamma = 0.85$, which implies a profit share of 0.15 in output, roughly the long-run average value of the profit share in Barkai and Benzell (2024). We set $\beta = \gamma \times 2/3$, which gives a labor share of 0.57, in line with the BEA's (0.55) and the BLS's (0.60) estimates for our sample period. The resulting capital share is $\alpha = \gamma - \beta = 0.28$. In Section 5, we show that alternative calibrations that allow input shares to vary across sectors or over time yield similar results.

4 Results

Our primary finding is that eliminating tax heterogeneity would improve TFP in the U.S. Our baseline estimates put this gain between 4 and 5 percent. When we decompose the potential TFP gains from eliminating tax heterogeneity into its three components, we find that the positive correlation of tax rates with capital and especially labor distortions accounts for the largest share of the TFP gains. The pure dispersion in tax rates explains a sizable part of the estimated TFP gains, but this component became smaller after the tax reforms of the 1980s.

Figure 2 shows the estimated TFP gains associated with eliminating tax heterogeneity. The solid black line shows our baseline result. The gain is positive throughout our sample period because, on average, factors are more productive at firms that face higher marginal tax rates. This suggests that tax heterogeneity exacerbates capital and labor distortions. The TFP gain ranges from 4 to 5 percent during the early 1980s to around 3 percent in recent years. The average gain across all years is 4.1%, which corresponds to four years of average annual U.S. TFP growth during our sample period.

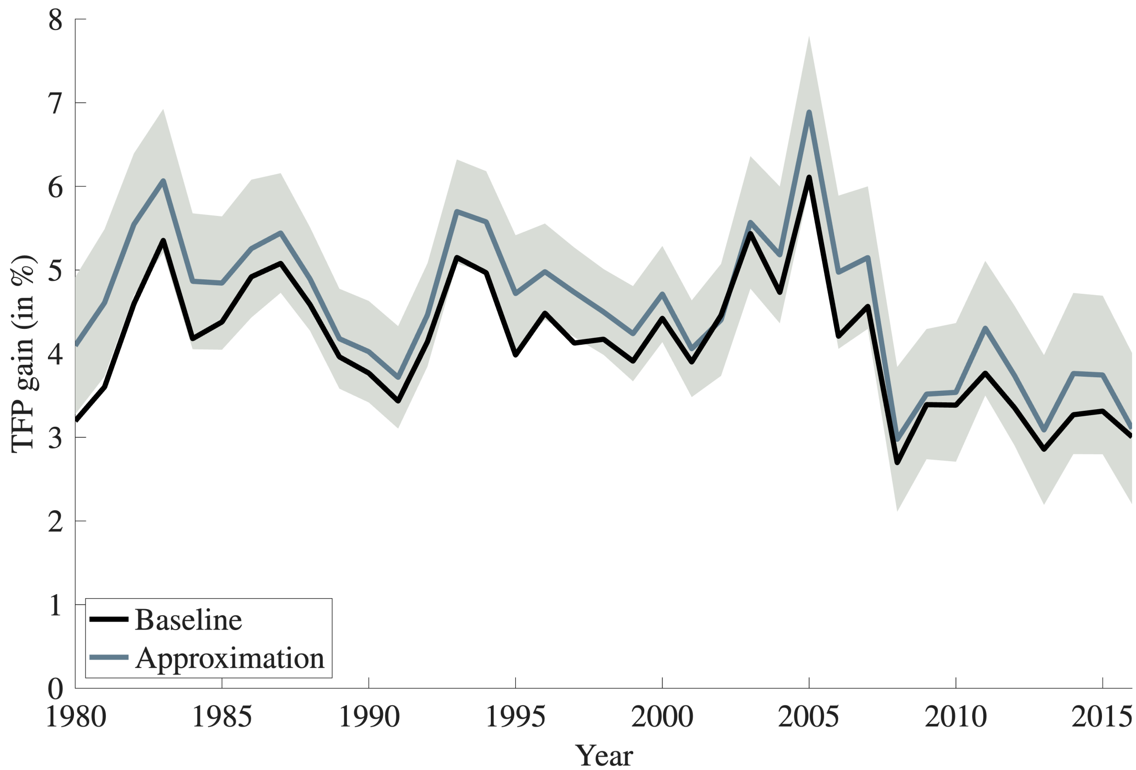


Figure 2: TFP gains from tax rate equalization

Note.– The figure shows the estimated change in aggregate TFP if all firms were to face a common corporate tax rate. Baseline model and its approximation using equation (7).

To understand the sources of the productivity gain, we use our approximation formula in equation (7). The gray line in Figure 2 shows the approximate TFP gains and the shaded gray area represents error bands corresponding to two standard errors above and below the point estimate. Using the approximation, the average TFP gain is 4.6%, close to the 4.1% benchmark. This reinforces our main finding that tax heterogeneity reduces aggregate TFP. The two series are highly correlated ($\rho = 0.96$), lending credence to our decomposition results, which we turn to next.

From equation (7), the total gains can be split into three parts, which are shown in Figure 3. The total productivity gains are shown as the solid gray line in Figure 3, which is identical to the gray line in Figure 2. The blue bars reflect the misallocation from pure tax heterogeneity. If tax wedges were uncorrelated with the marginal products of capital and labor, the blue bars would represent the total productivity gain from tax asymmetries in an otherwise frictionless economy. This component was 1-2% prior to 1986, mainly due to the higher statutory corporate tax rate, and has been under 1% since then, amounting to less than a quarter of the total TFP gain.

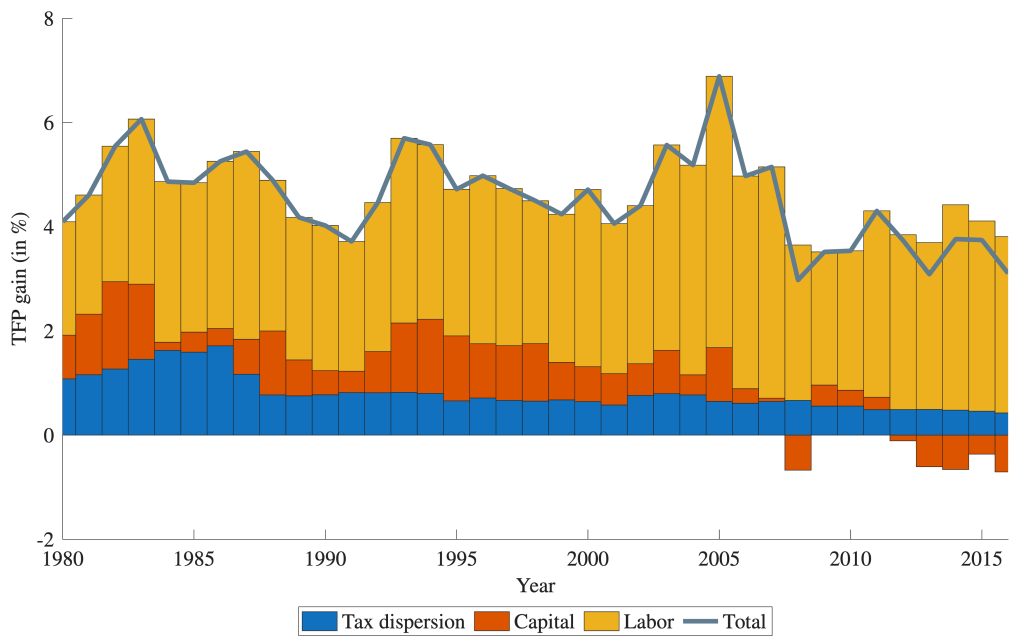


Figure 3: Components of TFP gains from eliminating tax heterogeneity

Note.– The figure shows the three components of TFP gains from tax rate equalization across firms (see equation (7)). The blue bars labeled "Tax dispersion" represent gains in an otherwise frictionless economy. The red bars labeled "Capital" show additional gains/losses in an economy with capital distortions. The bars labeled "Labor" show additional gains in an economy with capital and labor distortions.

The red bars represent the TFP gains due to the correlation between tax wedges and capital distortions. In most years, tax rates have a positive but weak correlation with capital distortions, resulting in additional TFP gains of less than 1%. The potential gains implied by this component have diminished over time, turning negative in recent years. This suggests an efficiency improvement from recent changes in the distribution of tax rates.

The yellow bars represent the TFP gain associated with the correlation between labor distortions and tax wedges. When that correlation is positive, equalizing taxes reduces rates for firms with higher marginal product of labor. Because of capital-labor complementarity, the ensuing employment reallocation toward those firms raises TFP. Quantitatively, this is the largest component of the TFP gain, partly because labor's share of income is roughly twice the capital share.[14]

Overall, the majority of the estimated TFP gain is attributable to the positive correlation between tax wedges and labor distortions. This highlights the importance of taking other distortions into account when studying the allocative effects of a particular distortion.

5 Robustness, extensions and long-run implications

Our baseline analysis quantifies the productivity losses from corporate tax heterogeneity under relatively conservative assumptions about firm behavior and the macroeconomic setting. In this section, we explore the robustness of these estimates, examine how they change under alternative modeling assumptions, and discuss their implications for long-run output. Table 2 summarizes our results. Appendix C provides the detailed methodology for different modeling extensions and presents the full results for each of the subsections below.

5.1 Long-run output effects

The productivity gains from tax homogenization translate into larger long-run output effects through induced capital accumulation. In a standard growth model with inelastic labor supply, our base-

[14] More generally, capital and labor co-move during the reallocation from eliminating tax heterogeneity as long as the elasticity of substitution between capital and labor is not too high (less than $1/(1-\gamma)$).

line 4.1% TFP gain implies a 5.7% increase in long-run output per capita, as higher productivity encourages additional investment. Models incorporating elastic labor supply would predict even larger effects.

A related consequence of output growth is higher tax revenue. Revenue-neutral tax reforms call for a lower tax rate in that case, which further encourages capital accumulation. This generates a fiscal multiplier effect: productivity-driven output growth increases tax revenue, permitting lower average tax rates that further stimulate capital formation. Accounting for this feedback loop, we estimate revenue-neutral tax homogenization would raise long-run output per capita by approximately 6.3%.

5.2 Interactions between distortions

Our baseline model treats the distribution of non-tax distortions, ω_R and ω_N, as exogenous to tax policy. When corporate taxation interacts with other economic frictions, the welfare effects of tax reform may exceed our baseline estimates. The results are shown under the heading 'Structural Distortions' in Table 2.

We consider two channels through which such interactions could amplify productivity gains. First, if labor costs are imperfectly deductible—whether due to incomplete expensing provisions or cash-in-advance requirements—corporate tax heterogeneity directly distorts employment decisions beyond its effects on capital allocation. Second, if the observed correlation between tax rates and capital costs reflects a causal relationship (for instance, through collateral constraints that bind more tightly at higher tax rates), then equalizing tax rates would simultaneously reduce dispersion in capital costs. Under the extreme assumption that all observed correlations between taxes and factor costs are structural, we estimate that productivity gains could reach an upper bound of 10.7 percent if tax homogenization mitigates underlying labor distortions and 5.7 percent if it mitigates capital distortions (see Appendix Section C.2 for the theoretical framework and detailed calculations). We interpret these scenarios as providing upper bounds on potential gains, as the true extent of such interactions remains model-specific and empirically ambiguous.

	Total	Capital-Linked	Labor-Linked	Tax Alone
Benchmark	4.6	0.6	3.2	0.8
Structural Distortions				
Capital	5.7	0.8	4.2	0.8
Labor	10.7	1.3	3.8	5.6
Declining labor share and				
rising capital share	4.7	0.6	3.3	0.9
rising markup	4.0	0.5	2.7	0.8

Table 2: TFP Gains from Homogenizing Tax Rates: Extensions and Robustness

Note.– Percent change in aggregate TFP if all firms were to face a common corporate tax rate. Average estimates over the 1980 - 2021 period. Benchmark model assumes corporate taxes only distort investment. The middle panel allows taxes to structurally affect pretax capital or labor costs. Bottom panel allows for declining labor share of income compensated by either higher capital share or higher price markups.

5.3 Robustness to macroeconomic trends

Our baseline estimate uses stable factor income shares throughout the sample period. The labor share of income declined from 57.7% to 53.0% between 1980 and 2016, raising questions about whether this trend affects our findings. In Appendix C.3, we evaluate two scenarios: one attributing the labor share decline entirely to rising capital intensity, and another attributing it to increasing markups. The results are summarized in the last two rows of Table 2.

Under rising capital shares, TFP gains remain nearly identical to the baseline, as the increased importance of capital in production roughly offsets compositional effects. Under rising markups, the estimated gains decline modestly—by approximately one percentage point in recent years—because higher markups move the economy away from the perfectly competitive benchmark where reallocation gains are maximized. Neither scenario substantially alters our central conclusions about the magnitude of productivity losses from tax heterogeneity.

5.4 Sectoral decomposition of productivity gains

Our baseline model estimates aggregate productivity gains using a common, economy-wide labor share parameter for all sectors. In Appendix C.4, we re-estimate TFP gains allowing for sector-specific factor elasticities that reflect technological differences across industries. This refinement yields a total TFP gain of 4.3%, similar to our benchmark estimate of 4.1%, confirming that our main findings are robust to technological heterogeneity across sectors. Decomposing the sources of these gains reveals that the vast majority—approximately 3.3 percentage points, or 77% of the total effect—derives from an improved allocation of capital and labor within sectors. The reallocation of productive factors between sectors contributes an additional 1.0 percentage point to aggregate productivity.

6 Conclusion

Our findings show that policies that seek to reduce differences in marginal corporate income tax rates would generate significant gains in aggregate productivity and long-run output. The majority of these gains are attributable to the empirical correlation between labor productivity and marginal tax rates: firms that face higher tax rates are typically those where labor is more productive. This highlights the importance of accounting for other frictions when studying the implications of a specific distortion to aggregate efficiency.

It is worth noting that although we estimate TFP losses associated with tax heterogeneity, eliminating the heterogeneity altogether is not necessarily productivity maximizing in distorted economies. Tax designs that seek to maximize aggregate productivity would aim to offset other distortions through firm-specific tax rates. The efficient design of corporate taxation is a promising avenue for future research.

References

ADAMOPOULOS, T., L. BRANDT, J. LEIGHT, AND D. RESTUCCIA (2022): "Misallocation, Selection, and Productivity: A Quantitative Analysis With Panel Data From China," *Econometrica*, 90, 1261–1282.

AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109, 659–684.

AUERBACH, A. J. (1983): "Taxation, Corporate Financial Policy and the Cost of Capital," *Journal of Economic Literature*, 21, 905–940.

——— (1986): "The Dynamic Effects of Tax Law Asymmetries," *The Review of Economic Studies*, 53, 205–225.

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

BARKAI, S. AND S. G. BENZELL (2024): "70 years of US corporate profits," *Journal of Corporate Finance*, 87, 102622.

BILS, M., P. J. KLENOW, AND C. RUANE (2021): "Misallocation or Mismeasurement?" *Journal of Monetary Economics*, 124, S39–S56.

BLOUIN, J., J. E. CORE, AND W. GUAY (2010): "Have the tax benefits of debt been overestimated?" *Journal of Financial Economics*, 98, 195–213.

CUMMINS, J. G., K. HASSETT, AND R. G. HUBBARD (1994): "A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments," *Brookings Papers on Economic Activity*, 25, 1–74.

DAVID, J. M. AND V. VENKATESWARAN (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109, 2531–67.

DJANKOV, S., T. GANSER, C. MCLIESH, R. RAMALHO, AND A. SHLEIFER (2010): "The Effect of Corporate Taxes on Investment and Entrepreneurship," *American Economic Journal: Macroeconomics*, 2, 31–64.

GLOVER, B. AND O. LEVINE (2023): "Corporate Tax Avoidance, Firm Size, and Capital Misallocation," SSRN Working Paper No. 4699638. Available at https://ssrn.com/abstract=4699638.

GOPINATH, G., Ş. KALEMLI-ÖZCAN, L. KARABARBOUNIS, AND C. VILLEGAS-SANCHEZ (2017): "Capital Allocation and Productivity in South Europe," *The Quarterly Journal of Economics*, 132, 1915–1967.

GRAHAM, J. R. (1996a): "Debt and the marginal tax rate," *Journal of financial Economics*, 41, 41–73.

——— (1996b): "Proxies for the corporate marginal tax rate," *Journal of financial economics*, 42, 187–221.

GRAHAM, J. R. AND L. F. MILLS (2008): "Using tax return data to simulate corporate marginal tax rates," *Journal of Accounting and Economics*, 46, 366–388.

GRAVELLE, J. G. (2001): "Whither Tax Depreciation?" *National Tax Journal*, 54, 513–526.

HALL, R. E. AND D. W. JORGENSON (1967): "Tax Policy and Investment Behavior," *The American Economic Review*, 57, 391–414.

HARBERGER, A. C. (1962): "The Incidence of the Corporation Income Tax," *Journal of Political Economy*, 70, 215–240.

——— (1964): "The Measurement of Waste," *The American Economic Review*, 54, 58–76.

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

KAYMAK, B. AND I. SCHOTT (2019): "Loss-offset provisions in the corporate tax code and misallocation of capital," *Journal of Monetary Economics*, 105, 1–20.

KIM, S. (2023): "Optimal Carbon Taxes and Misallocation across Heterogeneous Firm," Tech. rep., University of Maryland.

KLENOW, P., E. PASTÉN, AND C. RUANE (2025): "Carbon Taxes and Misallocation in Chile," Working paper.

LIPSEY, R. G. AND K. LANCASTER (1956): "The General Theory of Second Best," *The Review of Economic Studies*, 24, 11–32.

MIDRIGAN, V. AND D. Y. XU (2014): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104, 422–58.

QIAN, L. AND G. VERESHCHAGINA (2024): "Capital financing constraints, size-dependent distortions, and aggregate productivity," Working paper.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 11, 707–720.

SLATTERY, C. AND O. ZIDAR (2020): "Evaluating State and Local Business Incentives," *Journal of Economic Perspectives*, 34, 90–118.

STANDARD & POOR'S (2023): "Compustat Database. https://wrds-www.wharton.upenn.edu, Data retrieved from Wharton Research Data Services (WRDS) on 2023-09-01. © 2025. All rights reserved. Used with permission." .

STEFANSKI, R., L. AHLVIK, J. J. ANDERSEN, T. HARDING, AND A. TREW (2025): "Extracting Wedges: Misallocation and Taxation in the Oil Industry," CESifo Working Paper Series 11915, CESifo.

A Theoretical Appendix

This appendix details the derivations in the text and gives the formal proofs for the propositions.

A.1 Reformulating the model with an effective tax rate

The expression for the tax wedge τ^i_ω in (3) can also be derived from a setup in which the firm faces a *pre-tax* cost of capital of ω^i_R and a single effective tax rate τ^i on operating surplus, which includes loss offsets, as well as depreciation allowances:

$$\max_{k,n} \, (1-\tau^i)\left[zF(k,n) - \omega^i_N n\right] - \omega^i_R k, \tag{A1}$$

The first-order condition of the firm problem in (A1) is

$$zF_k(k,n) = w^i_R \cdot \frac{1}{1-\tau^i}, \tag{A2}$$

where $\omega^i_R = r^i + \delta^i$ are pre-tax user-costs. The conditions in (A2) and (3) are equivalent if we define

$$\tau^i = 1 - \frac{1-\tau^i_b}{1-\tau^i_b\lambda^i}. \tag{A3}$$

A.2 Effective tax rate with debt financing

We now consider a more general problem where a fraction of the marginal investment is financed with debt. A fraction ϕ^i of the marginal capital investment is debt-financed, while the remaining fraction $1-\phi^i$ is equity-financed (Auerbach, 1983). In that case, the first-order condition of the firm problem in (2) with respect to capital is:

$$zF_k(k,n) = \omega^i_R \cdot \underbrace{\frac{1-\tau^i_b\lambda^i + \phi\left[i^i(1-\tau^i_b) - r^i\right]}{1-\tau^i_b}}_{\omega^i_\tau}, \tag{A4}$$

where r^i is the after-tax return required by shareholders, and i^i is the interest rate paid to debt holders. With $\phi^i = 0$ we obtain the cost of financing capital under all-equity finance (as in (3)); with $\phi > 0$, the second term in the numerator adjusts the baseline cost by the difference between the after-tax cost of debt and the required return on equity.

As in Appendix A.1, we can derive an implied effective tax rate

$$\tau^i = 1 - (1-\tau_b^i)\frac{r^i + \delta^i}{(r^i+\delta^i)(1-\tau_b^i\lambda^i) + \phi^i\left[i(1-\tau_b^i) - r^i\right]}, \tag{A5}$$

which captures all the information about the loss offsets, depreciation allowances λ^i, and the financing mix. This lets us write the problem as in (A1).

A.3 An investment problem with multiple distortions

To fix ideas, we develop an investment problem with multiple distortions. For notational simplicity, we model only the effective (post-deduction) marginal tax rate τ, as derived in Appendices A.1 and A.2. Consider the investment problem of a firm that uses capital and labor to produce output with the decreasing-returns-to-scale (DRS) production function, $y = zF(k, n)$, using technology z. Each period, the firm chooses capital investment and labor to maximize the payout to shareholders, which depends on current and expected after-tax profits. For simplicity, assume that the firm knows the realization of next period's productivity z and taxes τ at the time of the investment decision. The firm's problem is given by

$$\begin{aligned} V(z,k) &= \max_{n',k',i} -i - \Phi\left(\frac{i}{k}\right)k \\ &+ \beta\left[(1-\tau)\left[zF(k',n') - \omega_N n'\right] + E_{z'|z}V(z',k')\right] \end{aligned} \tag{A6}$$

subject to the law of motion for capital and a collateral constraint

$$k' = i + (1-\delta)k \tag{A7}$$

$$i \leq \zeta q k, \tag{A8}$$

with associated multipliers q and μ. The first-order condition for labor is

$$zF_n(k,n) = \omega_N, \tag{A9}$$

where ω_N denotes the effective user cost of labor. The optimality conditions for investment and future capital of the firm's problem in are given by the first-order condition with respect to investment:

$$i: \quad -1 - \Phi'\left(\frac{i}{k}\right) + q - \mu = 0 \quad \Leftrightarrow \quad q = 1 + \Phi'\left(\frac{i}{k}\right) + \mu. \tag{A10}$$

and

$$k': \quad \beta\left[(1-\tau)zF_k(k',n') + E_{z'|z}V_k(z',k')\right] - q = 0, \tag{A11}$$

with the envelope condition

$$V_k(z,k) = -\Phi\left(\frac{i}{k}\right) + \Phi'\left(\frac{i}{k}\right)\frac{i}{k} + q(1-\delta) + \mu\zeta q. \tag{A12}$$

This implies the following optimality condition for the choice of capital:

$$q = \beta\left[(1-\tau)zF_k(k',n') - \Phi\left(\frac{i'}{k'}\right) + \Phi'\left(\frac{i'}{k'}\right)\frac{i'}{k'} + q'(1-\delta) + \mu'\zeta' q'\right] \tag{A13}$$

The marginal cost of capital, q, is balanced against the discounted future benefit of an additional unit of capital. This benefit consists of the marginal product of capital (net of taxes), the value of non-depreciated capital, and the effects on the marginal costs of investment as well as on the financial constraint. To simplify , define the tax wedge ω_τ as $\omega_\tau \equiv \frac{1}{1-\tau}$, and the remainder as

$$\omega_R \equiv \frac{q}{\beta} - q'(1-\delta) - \mu'\zeta'q' + E_{z'|z}\left[\Phi\left(\frac{i'}{k'}\right) + \Phi'\left(\frac{i'}{k'}\right)\frac{i'}{k'}\right]. \tag{A14}$$

Using the definitions above, we can write as

$$zF_k(k,n) \quad = \quad \omega_K \quad \equiv \quad \omega_\tau \cdot \omega_R, \tag{A15}$$

where ω_K denotes the effective user cost of capital, comprising tax costs, ω_τ, and costs of capital adjustment or financial constraints, summarized by ω_R. Any additional frictions a firm might be facing would be included in ω_R.

A.4 Proofs and derivations

Proposition 1. *TFP in the distorted economy is*

$$Z = \frac{Y}{K^\alpha N^\beta} = \frac{\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{\beta}{1-\gamma}} {\omega_K}^{-\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{\beta}{1-\gamma}} {\omega_K}^{-\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{1-\alpha}{1-\gamma}} {\omega_K}^{-\frac{\alpha}{1-\gamma}} dG\right]^\beta}. \tag{4}$$

Proof. The profit-maximizing levels of capital, labor, and output are:

$$n = z^{\frac{1}{1-\gamma}} \left(\frac{\beta}{\omega_N}\right)^{\frac{1-\alpha}{1-\gamma}} \left(\frac{\alpha}{\omega_K}\right)^{\frac{\alpha}{1-\gamma}} \tag{A16}$$

$$k = z^{\frac{1}{1-\gamma}} \left(\frac{\beta}{\omega_N}\right)^{\frac{\beta}{1-\gamma}} \left(\frac{\alpha}{\omega_K}\right)^{\frac{1-\beta}{1-\gamma}} \tag{A17}$$

$$y = z^{\frac{1}{1-\gamma}} \left(\frac{\beta}{\omega_N}\right)^{\frac{\beta}{1-\gamma}} \left(\frac{\alpha}{\omega_K}\right)^{\frac{\alpha}{1-\gamma}} \tag{A18}$$

Substituting in the definition of TFP gives the following:

$$Z = \frac{Y}{K^\alpha N^\beta} = \frac{\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{\beta}{1-\gamma}} {\omega_K}^{-\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{\beta}{1-\gamma}} {\omega_K}^{-\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{1-\alpha}{1-\gamma}} {\omega_K}^{-\frac{\alpha}{1-\gamma}} dG\right]^\beta}$$

□

Proposition 2. *TFP in the counterfactual economy with homogeneous tax rates across firms is*

$$Z^* = \frac{\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{\beta}{1-\gamma}} {\omega_K}^{-\frac{\alpha}{1-\gamma}} {\omega_\tau}^{\frac{\alpha}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{\beta}{1-\gamma}} {\omega_K}^{-\frac{1-\beta}{1-\gamma}} {\omega_\tau}^{\frac{1-\beta}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} {\omega_N}^{-\frac{1-\alpha}{1-\gamma}} {\omega_K}^{-\frac{\alpha}{1-\gamma}} {\omega_\tau}^{\frac{\alpha}{1-\gamma}} dG\right]^\beta}. \tag{5}$$

Proof. In the counterfactual economy, ${\omega_K}' = \bar{\omega}^K \omega_R = \bar{\omega}^K \omega_K / \omega_\tau$ and ${\omega_N}' = \bar{\omega}_n \omega_N$. Proof follows from substituting these in equation (4). The common components of distortions, $\bar{\omega}_n$ and $\bar{\omega}^K$, cancel each other out. □

Proposition 3. *If ω_K, ω_N, and ω_τ are jointly log-normally distributed, then eliminating tax heterogeneity yields the following change in aggregate TFP:*

$$\ln \frac{Z^*}{Z} = \frac{1}{2} \frac{\alpha(1-\beta)}{1-\gamma} \sigma_\tau^2 + \frac{\alpha(1-\beta)}{1-\gamma} \sigma_\tau^2 (L_{k\tau} - 1) + \frac{\alpha\beta}{1-\gamma} \sigma_\tau^2 L_{n\tau}, \tag{7}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_K$ and $\ln \omega_N$ on $\ln \omega_\tau$.

Proof. Let $\mu_x = \mathbb{E}[\ln x]$ and $\sigma_x^2 = \mathbb{V}[\ln x]$ be the mean and variance of the log of a variable x. Under joint log-normality:

$$\begin{aligned} \ln \int y \quad = \quad & \frac{\alpha}{1-\gamma} \ln \alpha + \frac{\beta}{1-\gamma} \ln \beta + \frac{1}{1-\gamma} (\mu_z - \alpha\mu_k - \beta\mu_n) \\ & + \frac{1}{2} \frac{1}{(1-\gamma)^2} (\sigma_z^2 + \alpha^2 \sigma_k^2 + \beta^2 \sigma_n^2 - 2\alpha\sigma_{zk} - 2\beta\sigma_{zn} + 2\alpha\beta\sigma_{kn}) \end{aligned}$$

$$\begin{aligned}\ln \int k \;&=\; \frac{(1-\beta)}{1-\gamma}\ln\alpha + \frac{\beta}{1-\gamma}\ln\beta + \frac{1}{1-\gamma}(\mu_z - (1-\beta)\mu_k - \beta\mu_n) \\ &+\frac{1}{2}\frac{1}{(1-\gamma)^2}(\sigma_z^2 + (1-\beta)^2\sigma_k^2 + \beta^2\sigma_n^2 - 2(1-\beta)\sigma_{zk} - 2\beta\sigma_{zn} + 2(1-\beta)\beta\sigma_{kn})\end{aligned}$$

$$\begin{aligned}\ln \int n \;&=\; \frac{\alpha}{1-\gamma}\ln\alpha + \frac{(1-\alpha)}{1-\gamma}\ln\beta + \frac{1}{1-\gamma}(\mu_z - \alpha\mu_k - (1-\alpha)\mu_n) \\ &+\frac{1}{2}\frac{1}{(1-\gamma)^2}(\sigma_z^2 + \alpha^2\sigma_k^2 + (1-\alpha)^2\sigma_n^2 - 2\alpha\sigma_{zk} - 2(1-\alpha)\sigma_{zn} + 2\alpha(1-\alpha)\sigma_{kn})\end{aligned}$$

Using these equations, the TFP in the distorted economy is:

$$\ln Z = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_k^2 - \beta(1-\alpha)\sigma_n^2 - 2\alpha\beta\sigma_{kn}\right]. \tag{A19}$$

When tax differentials are eliminated capital distortions are given simply by ω_R, which gives the TFP in that counterfactual as:

$$\ln Z^* = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_R^2 - \beta(1-\alpha)\sigma_n^2 - 2\alpha\beta\sigma_{Rn}\right]. \tag{A20}$$

Note that $\sigma_k^2 - \sigma_R^2 = \sigma_\tau^2 + 2\sigma_{R\tau}$ and $\sigma_{kn} - \sigma_{Rn} = \sigma_{n\tau}$. Rearranging the terms and netting out μ_z,

σ_z^2 and σ_n^2 terms, the efficiency losses from distortions are equivalent to:

$$
\begin{aligned}
\ln \frac{Z^*}{Z} &= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_k^2-\sigma_R^2)+2\alpha\beta(\sigma_{kn}-\sigma_{Rn})\right] \\
&= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_\tau^2+2\sigma_{R\tau})+2\alpha\beta\sigma_{n\tau}\right] \\
&= \frac{\sigma_\tau^2}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(1+2\frac{\sigma_{R\tau}}{\sigma_\tau^2})+2\alpha\beta\frac{\sigma_{n\tau}}{\sigma_\tau^2}\right] \\
&= \frac{\sigma_\tau^2}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(1+2(L_{k\tau}-1))+2\alpha\beta L_{n\tau}\right] \\
&= \frac{1}{2}\frac{1}{1-\gamma}\alpha(1-\beta)\sigma_\tau^2+\frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1)+\frac{\alpha\beta}{1-\gamma}\sigma_\tau^2 L_{n\tau}
\end{aligned}
$$

The last two equalities substitute the linear projection coefficients for $\frac{\sigma_{R\tau}}{\sigma_\tau^2}=L_{R\tau}=L_{k\tau}-1$ and $\frac{\sigma_{n\tau}}{\sigma_\tau^2}=L_{n\tau}$. □

Proposition 4. *Assume that the tax wedge is measured with error, $\omega_\tau^*=\omega_\tau+\epsilon$ with $\mathbb{E}(\epsilon|\omega_\tau)=0$. Then, replacing ω_τ by ω_τ^* in equation (7) underestimates the net TFP gain from eliminating tax heterogeneity:*

$$\ln(Z^*/Z)|\omega_\tau^*=\ln(Z^*/Z)|\omega_\tau-\frac{\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}. \tag{8}$$

Proof. Let $\sigma_{\tau^*}^2=\sigma_\tau^2+\sigma_\epsilon^2$ denote the variance of the measured tax wedge. Define the projection $x=L_{x\tau}\times\ln\omega_\tau+e_x$, where ω_τ is the true tax wedge. The OLS estimate of $L_{x\tau}$ from the projection of x on $\ln\omega_{\tau^*}^2$ is: $\hat{L}_{x\tau}^{OLS}=L_{x\tau}\times\frac{\sigma_\tau^2}{\sigma_{\tau^*}^2}$.

$$
\begin{aligned}
\ln(Z^*/Z)|\omega_\tau^* &= \frac{\alpha(1-\beta)}{1-\gamma}\frac{\sigma_{\tau^*}^2}{2}+\frac{\alpha(1-\beta)}{1-\gamma}\frac{\sigma_{\tau^*}^2}{2}(\hat{L}_{k\tau}^{OLS}-1)+\frac{\alpha\beta}{1-\gamma}\frac{\sigma_{\tau^*}^2}{2}\hat{L}_{n\tau}^{OLS} \\
&= -\frac{\sigma_\tau^2+\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}+\frac{\alpha(1-\beta)}{1-\gamma}\frac{\sigma_\tau^2}{2}L_{k\tau}+\frac{\alpha\beta}{1-\gamma}\frac{\sigma_\tau^2}{2}\hat{L}_{n\tau} \\
&= \ln(Z^*/Z)|\omega_\tau-\frac{\sigma_\epsilon^2}{2}\frac{\alpha(1-\beta)}{1-\gamma}.
\end{aligned}
$$

□

A.5 Computation of the non-linear gains and measurement error correction

For each $x \in \{\omega_K, \omega_N\ z\}$, define $\ln \hat{x} = b_{x\tau} \times \ln \omega_\tau$, where $b_{x\tau}$ is the OLS estimator of $L_{x\tau}$. We ignore the intercept term in that projection, because it cancels out from the TFP equations as it is common to all firms. Note also that we are using the OLS estimator, not the consistent estimator.[A1]

We then compute current TFP by substituting $\hat{\omega}_K$, $\hat{\omega}_N$ and $\hat{z}$ in equation (4), and the counterfactual TFP by substituting $\hat{\omega}_N$, $\hat{z}$, and $\omega'_k = \hat{\omega}_K/\omega_{\tau^*}$ for ω_K in the same equation. This yields the following equations:

$$\hat{Z}_{\tau^*} = \frac{\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-\alpha b_{k\tau})} dG_{\tau^*}}{\left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-(1-\beta) b_{k\tau})} dG_{\tau^*}\right]^{\alpha} \left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-(1-\alpha) b_{n\tau}-\alpha b_{k\tau})} dG_{\tau^*}\right]^{\beta}}$$

$$\hat{Z}^*_{\tau^*} = \frac{\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-\alpha b_{R\tau})} dG_{\tau^*}}{\left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-\beta b_{n\tau}-(1-\beta) b_{R\tau})} dG_{\tau^*}\right]^{\alpha} \left[\int (\omega_{\tau^*})^{\frac{1}{1-\gamma}(b_{z\tau}-(1-\alpha) b_{n\tau}-\alpha b_{R\tau})} dG_{\tau^*}\right]^{\beta}},$$

where $b_{R\tau} = b_{k\tau} - 1$ is the projection coefficient of $\ln \omega_R$ on $\ln \omega_{\tau^*}$ and G_{τ^*} is the marginal distribution of the measured tax wedge.

When ϵ is distributed independently log-normal, then for any scalar $c > 0$, $E[\omega^c_{\tau^*}] = E[\omega^c_\tau \epsilon^c] = E[\omega^c_\tau] \cdot E[\epsilon^c] = E[\omega^c_\tau] \cdot \exp(c^2\sigma^2_\epsilon/2)$. Replacing c with the appropriate power component for each term gives:

$$\ln \frac{\hat{Z}^*_{\tau^*}}{\hat{Z}_{\tau^*}} = \ln \frac{\hat{Z}^*_{\tau}}{\hat{Z}_{\tau}} - \frac{\sigma^2_\epsilon}{2} \frac{\alpha(1-\beta)}{1-\gamma}. \tag{A21}$$

Repeating these steps for the alternative formulations in Section C.2 by substituting τ^* for τ in the expressions for $\hat{Z}^{**}_{\tau^*}$ and $\hat{Z}^{***}_{\tau^*}$ presented in equations (A25) and (A27) give the following relations between the measured TFP loss and the actual TFP loss:

[A1] Because various biases from using the OLS estimator cancel each other out, the resulting formula for error correction is simpler and aligns with that reported in Proposition 4. We also considered using the consistent (IV) estimator, and a non-parametric estimator for $\mathbb{E}[\ln x|\omega_t au]$. Both yield similar results.

$$\ln \frac{\hat{Z}_{\tau^*}^{**}}{\hat{Z}_{\tau^*}} = \ln \frac{\hat{Z}_{\tau}^{**}}{\hat{Z}_{\tau}} - \frac{\sigma_\epsilon^2}{2} \frac{\gamma}{1-\gamma}. \tag{A22}$$

$$\ln \frac{\hat{Z}_{\tau^*}^{***}}{\hat{Z}_{\tau^*}} = \ln \frac{\hat{Z}_{\tau}^{***}}{\hat{Z}_{\tau}} \times \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \sigma_\tau^2}. \tag{A23}$$

B Data Appendix

The firm-level data used in Section 3 were constructed as follows. We use the annual Compustat database, (Standard & Poor's, 2023), which provides balance-sheet data on publicly listed companies in the U.S. Our sample includes the years 1980–2021. We restrict attention to firms registered in the U.S. and exclude firms in the finance, insurance, and real estate sectors, as well as in utilities and public administration. We remove observations with negative sales.

We construct firm-level capital stocks by using a perpetual inventory method. For each firm, we start with the first year in which information on gross and net property, plant, and equipment (PPEGT and PPENT) is available. We then build the capital stock by adding the change in PPENT deflated by the investment price deflator to the calculated capital stock for that year. Output y_{it} is the sum of sales and changes in inventories during the year. Firm-level TFP is computed as $\log z_{it} = \log y_{it} - \beta \cdot \log n_{it} - \alpha \cdot \log k_{it}$, where factor shares come from the BLS and are computed for the year 2000, the midpoint of our sample. We remove NAICS-year fixed effects from firm-level TFP.

We supplement these data with information on firms' marginal tax rates taken from *i)* Graham and Mills (2008), abbreviated as "GM" and *ii)* Blouin et al. (2010), abbreviated as "BCG".[A2] While the GM database covers the years 1980–2021, the BCG data are only available from 1980–2016. We link the Compustat data to the marginal tax rate data via the firm identifier GVKEY. Finally, we remove firm-year observations for which both tax rates are missing.

This results in a sample of 185,203 unique firm-year observations, averaging about 4,600 firms

[A2] Each study contains two measures of the marginal tax rates: before and after interest deductions are applied. The marginal tax rates we use in this paper are after interest deductions.

per year. Marginal effective tax rates are available for 70.3 percent of our observations (90.2 percent for the BCG tax rates).

Table A1 summarizes the two EMTR measures. Rates estimated by Graham and Mills (2008) show more bunching at zero and at the top statutory rate, which has varied over the years. Rates estimated by Blouin et al. (2010) display a smoother distribution, with a higher average rate and a slightly smaller variance. The two measures are highly but imperfectly correlated ($\rho = 0.61$).

Table A1: Summary statistics of marginal tax rates

Variable	mean	sd.	p25	p50	p75	N
τ^{GM}	0.169	0.171	0.007	0.070	0.342	125,048
τ^{BCG}	0.240	0.136	0.108	0.283	0.342	159,247

Note.– The two effective tax variables are taken from Graham and Mills (2008) and Blouin et al. (2010). See Appendix D for variable definitions and sample selection.

C Additional results

This appendix gives more details on the results presented in Section 5 in the main text.

C.1 Long-run output

Aggregate productivity gains imply output gains for given quantities of inputs. As was shown in section 5.1, in the long-run, improved productivity leads to capital investment, which further raises output. In the standard growth model with representative households and inelastic labor supply, steady-state capital is given by $K = (\alpha Z/r)^{\frac{1}{1-\alpha}}$, where Z is aggregate TFP. A percentage increase in TFP raises capital by $1/(1-\alpha) = 1.4\%$ and output by $\alpha/(1-\alpha) = 0.4\%$ in the long-run beyond the direct increase from higher TFP. Consequently, the baseline 4.1% TFP gain from tax homogenization translates to a 5.7% ($= 4.1 \times 1.4$) long-run gain in output per capita.[A3]

[A3]More generally, long-run output growth depends on the price elasticities of factor supply. Models with positively-sloped capital supply schedules, such as Aiyagari (1994), would predict smaller gains, and those with positively sloped labor supply schedules would predict larger gains than the figures reported here.

Our results on the revenue-neutral output gains in Section 5.1 were derived as follows. Pretax corporate income is proportional to output in our model: $\Pi = (1-\gamma)Y$. Each percent decline in the tax rate raises the capital stock by $\frac{1}{1-\alpha}\frac{\tau}{1-\tau}$, and output by α times the capital stock.[A4] Additional output allows for even lower tax rates, and so on. The associated multiplier is $\left(1 - \frac{\alpha}{1-\alpha}\frac{\tau}{1-\tau}\right)^{-1}$, which is about 1.1 when $\alpha = 0.28$ and the (common) tax rate is 20 percent, roughly the average effective rate in our sample between the two measures. Therefore, a revenue-neutral tax homogenization reform would raise output per capita by 6.3% (5.7×1.1) in the long-run, accounting for capital accumulation.

C.2 Interactions between distortions

When corporate taxation interacts with other distortions, eliminating tax differentials would also alter the distribution of ω_R and ω_N, which we assumed to fixed in the main text. An extension is discussed in section 5.2, which we provide more details on here. In models with collateral constraints or cash-in-advance requirements, lower taxes can free up internal funds for investment or employment, easing such constraints. This would raise productivity beyond the above estimates when highly-taxed firms are also more productive.

Ascertaining the extent of these additional gains however requires a more definitive stance on the nature of the underlying frictions than we adopt here. If investment distortions manifest as firm-specific interest costs as opposed to collateral constraints, for instance, eliminating tax heterogeneity may not alleviate capital distortions. Some progress can be made, however, on the *potential* role of these alternative models by interpreting the empirical correlation between productivity and tax rates as entirely endogenous. As we demonstrate next, this provides an upper bound on the productivity gains from homogenizing tax rates given the empirical patterns.

[A4] Because $\omega_\tau = \frac{1}{1-\tau}$, the percent change in the tax wedge is related to the percent change in the tax rate as follows: $\frac{d\omega_\tau}{\omega_\tau} = -\frac{d\tau}{\tau} \times \frac{\tau}{1-\tau}$. Multiplying that with the cost elasticity of capital demand, $1/(1-\alpha) = 1.5$, gives the total effect on the capital stock.

Labor distortions

Basic theory suggests that corporate taxes should not distort employment, conditional on capital and technology, because labor expenses are tax-deductible (equation (2)). Corporate tax rates can nonetheless directly affect employment when labor costs cannot be fully deducted or when there are explicit employment incentives. Similarly, cash-in-advance requirements, where wages have to be paid before sales are realized, generate a tax wedge in the optimality condition for employment. Because those funds can otherwise be invested in capital, the opportunity cost of capital, ω_K, become part of labor costs. Below, we present a cash-in-advance model and illustrate how corporate tax rates can directly affect employment in addition to their indirect effect through investment.

Suppose a fraction $\lambda \in [0, 1]$ of labor costs are not tax-deductible. Denoting the total cost per employee by ω_L, the after-tax income is now given by $(1-\tau)[zF(k, n)-(1-\lambda)\omega_L n]-\lambda\omega_L-\omega_R k$, which gives the following labor demand condition:

$$zF_n(k, n) = (1 - \lambda + \lambda\omega_\tau)\omega_L = \omega_N \tag{A24}$$

When $\lambda > 0$, a higher corporate tax rate raises the marginal cost of labor. Importantly, heterogeneity in tax rates can generate a dispersion in marginal costs of labor ω_N across firms, even when all firms face the same total cost per employee ω_L.[A5] Unlike in our baseline model, eliminating tax heterogeneity in this economy reduces the dispersion in marginal cost of labor, thereby leading to larger gains in aggregate productivity.

In our associated calculations we consider the extreme case of $\lambda = 1$, which implies $\omega_N = \omega_L \cdot \omega_\tau$. This is tantamount to interpreting the correlation between labor productivity and tax rates as structural. The following proposition gives the counterfactual TFP from eliminating tax heterogeneity in that scenario:

Proposition 1. *In models with imperfect labor expensing, TFP in the counterfactual economy with*

[A5] A similar distortion can also arise when $\lambda < 0$, i.e., when there is a tax credit for employment. We focus on $\lambda > 0$ here given the positive empirical correlation between labor productivity and tax rates below.

homogeneous tax rates across firms is:

$$Z^{**} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{\beta}{1-\gamma}} \omega_K^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{\gamma}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{\beta}{1-\gamma}} \omega_K^{-\frac{1-\beta}{1-\gamma}} \omega_\tau^{\frac{1}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_N^{-\frac{1-\alpha}{1-\gamma}} \omega_K^{-\frac{\alpha}{1-\gamma}} \omega_\tau^{\frac{1}{1-\gamma}} dG\right]^\beta}. \tag{A25}$$

Dividing Z^{**} above with Z in equation (4) gives the relative aggregate changes in TFP from homogenizing tax rates across firms. When all distortions are distributed log-normally, the implied change in aggregate TFP can be simplified as follows.

Proposition 2. *Suppose $\omega_K = \omega_R \cdot \omega_\tau$, $\omega_N = \omega_l \cdot \omega_\tau$, and ω_τ are jointly log-normally distributed. Eliminating tax heterogeneity yields the following change in aggregate TFP in models with imperfect labor expensing:*

$$\ln \frac{Z^{**}}{Z} = \frac{1}{2}\frac{\gamma}{1-\gamma}\sigma_\tau^2 + \frac{\alpha}{1-\gamma}\sigma_\tau^2(L_{k\tau} - 1) + \frac{\beta}{1-\gamma}\sigma_\tau^2(L_{n\tau} - 1), \tag{A26}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_K$ and $\ln \omega_N$ on $\ln \omega_\tau$.

Comparing equations (7) and (A26) reveals the significance of a structural link between taxes and labor costs. The first component, attributable to tax asymmetry, is now larger because it includes the direct contribution of tax heterogeneity to the dispersion in ω_N. The second component captures the correlation between tax rates and other capital distortions. Because taxes now raise marginal costs of both capital and labor, they generate a positive correlation between the two. This exacerbates the distortionary effect of taxes (via σ_{kn} in equation (6)). The last component in (A26) can be smaller or larger than in equation (7) because it now captures the productivity implications of the correlation between the tax wedge, ω_τ, and *other* labor distortions, ω_L.

Capital distortions

Similarly, consider the empirical decomposition of capital distortions into a random component and a tax-related component: $\ln \omega_R = L_{R\tau} \ln \omega_\tau + \ln \omega_R'$, where $\ln \omega_R'$ is orthogonal to $\ln \omega_\tau$. If this empirical relationship is structural, then eliminating tax differences would alter $\ln \omega_R$ through $L_{R\tau}$, unlike our baseline model. Consequently, setting $\omega_K{}' = \omega_R{}'$ in the counterfactual economy delivers the following TFP:

Proposition 3. *TFP in the counterfactual economy with homogeneous tax rates across firms is*

$$Z^{***} = \frac{\int z^{\frac{1}{1-\gamma}} \omega_N{}^{-\frac{\beta}{1-\gamma}} \omega_K{}^{-\frac{\alpha}{1-\gamma}} \omega_\tau{}^{\frac{\alpha L_{k\tau}}{1-\gamma}} dG}{\left[\int z^{\frac{1}{1-\gamma}} \omega_N{}^{-\frac{\beta}{1-\gamma}} \omega_K{}^{-\frac{1-\beta}{1-\gamma}} \omega_\tau{}^{\frac{(1-\beta)L_{k\tau}}{1-\gamma}} dG\right]^\alpha \left[\int z^{\frac{1}{1-\gamma}} \omega_N{}^{-\frac{1-\alpha}{1-\gamma}} \omega_K{}^{-\frac{\alpha}{1-\gamma}} \omega_\tau{}^{\frac{\alpha L_{k\tau}}{1-\gamma}} dG\right]^\beta}. \tag{A27}$$

The approximate gains from tax homogenization are as follows.

Proposition 4. *If ω_K, ω_N, and ω_τ are jointly log-normally distributed, then eliminating tax heterogeneity yields the following change in aggregate TFP:*

$$\ln \frac{Z^{***}}{Z} = \frac{1}{2}\frac{\alpha(1-\beta)}{1-\gamma}\sigma_\tau^2 L_{k\tau}^2 + \frac{\alpha\beta}{1-\gamma}\sigma_\tau^2 L_{n\tau} L_{k\tau}, \tag{A28}$$

where σ_τ^2 is the variance of $\ln \omega_\tau$ and $L_{k\tau} = \sigma_{k\tau}/\sigma_\tau^2$ and $L_{n\tau} = \sigma_{n\tau}/\sigma_\tau^2$ denote the slope coefficients from a linear projection of $\ln \omega_K$ and $\ln \omega_N$ on $\ln \omega_\tau$.

Proof. Substituting $\omega_N = \omega_L$ and $\omega_K = \omega_R$ in equation (A19) for the TFP in the distorted economy gives:

$$\ln Z = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_k^2 - \beta(1-\alpha)\sigma_n^2 - 2\alpha\beta\sigma_{kn}\right]. \tag{A29}$$

When tax differentials are eliminated capital distortions are given simply by ω_R, which gives the TFP in that counterfactual as:

$$\ln Z^* = \mu_z + \frac{1}{2}\frac{1}{1-\gamma}\left[\sigma_z^2 - \alpha(1-\beta)\sigma_R^2 - \beta(1-\alpha)\sigma_l^2 - 2\alpha\beta\sigma_{Rl}\right]. \tag{A30}$$

Note that $\sigma_k^2 - \sigma_R^2 = \sigma_\tau^2 + 2\sigma_{R\tau}$, $\sigma_n^2 - \sigma^2 = \sigma_\tau^2 + 2\sigma_{l\tau}$ and $\sigma_{kn} - \sigma_{Rl} = \sigma_{R\tau} + \sigma_{l\tau} + \sigma_\tau^2$. Rearranging the terms and netting out μ_z and σ_z^2 terms, the efficiency gains or losses from distortions are equivalent to:

$$\begin{aligned}
\ln\frac{Z^*}{Z} &= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_k^2-\sigma_R^2)+\beta(1-\alpha)(\sigma_n^2-\sigma_l^2)+2\alpha\beta(\sigma_{kn}-\sigma_{Rl})\right] \\
&= \frac{1}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(\sigma_\tau^2+2\sigma_{R\tau})+\beta(1-\alpha)(\sigma_\tau^2+2\sigma_{l\tau})\cdots\right. \\
&\qquad\left.\cdots+2\alpha\beta(\sigma_{R\tau}+\sigma_{l\tau}+\sigma_\tau^2)\right] \\
&= \frac{\sigma_\tau^2}{2}\frac{1}{1-\gamma}\left[\alpha(1-\beta)(1+2\frac{\sigma_{R\tau}}{\sigma_\tau^2})+\beta(1-\alpha)(1+2\frac{\sigma_{l\tau}}{\sigma_\tau^2})\cdots\right. \\
&\qquad\left.\cdots+2\alpha\beta(1+\frac{\sigma_{R\tau}}{\sigma_\tau^2}+\frac{\sigma_{l\tau}}{\sigma_\tau^2})\right]. \\
&= \frac{1}{2}\frac{\gamma}{1-\gamma}\sigma_\tau^2+\frac{\alpha}{1-\gamma}\sigma_\tau^2(L_{k\tau}-1)+\frac{\beta}{1-\gamma}\sigma_\tau^2(L_{n\tau}-1)
\end{aligned}$$

The last equality substitutes the linear projection coefficients for $\frac{\sigma_{R\tau}}{\sigma_\tau^2}=L_{R\tau}=L_{k\tau}-1$ and $\frac{\sigma_{l\tau}}{\sigma_\tau^2}=L_{l\tau}=L_{n\tau}-1$, and $\gamma=\alpha+\beta$. □

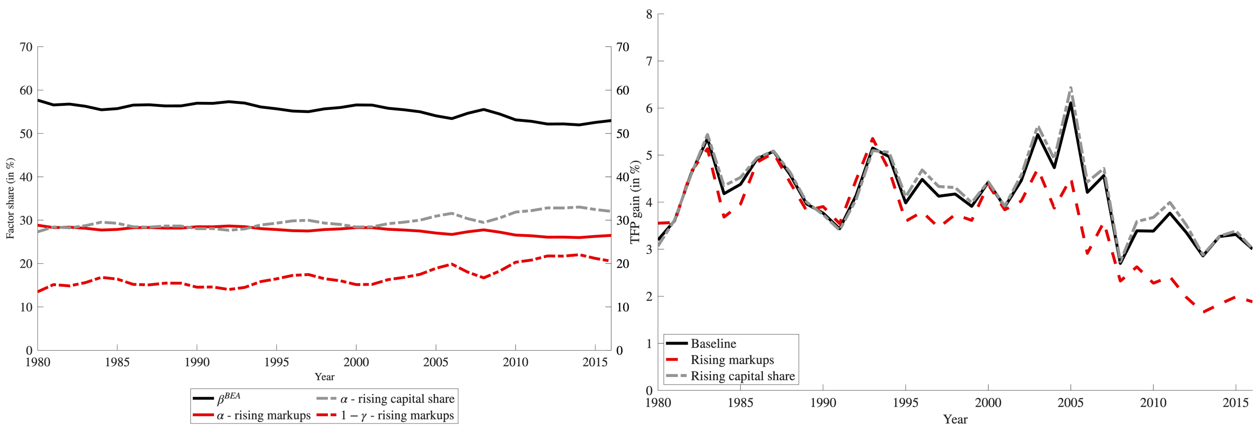
This gain is larger than the baseline if and only if $L_{R\tau}>0$, or, equivalently, if $L_{k\tau}>1$. Therefore, a structural interpretation of the empirical relationship between taxes and capital distortions amplifies the productivity effects of tax heterogeneity. When $L_{R\tau}>0$, eliminating tax differences also reduces the dispersion in ω_R, thereby alleviating capital distortions.

C.3 Macroeconomic trends and estimated TFP gains

Section 5.3 investigated how TFP gains from eliminating tax heterogeneity are affected by trends in factor shares and markups.

We take the decline of the labor share as given and benchmark its drop to the BEA's measure for each year of our analysis. Of course, the shares of capital, labor and profits sum to one, and a change in the labor share necessarily changes at least one other parameter. While the evidence on the decline in the aggregate labor share is relatively well-accepted, how much of that decline was redistributed to capital versus profits is less clear. We therefore consider two alternative scenarios. First, we assume that the decline in the labor share was matched one for one by a rise in the capital share of income, keeping the share of profits constant over time. Second, we assume that the labor share and capital declined proportionally, thus raising the profit share $1-\gamma$ over time.

The resulting values for α, β, and γ are shown in Panel (a) of Figure A1. The black solid line shows the labor share of income published by the BEA. It declines from 57.7 percent in 1980 to 53.0 percent in 2016. In the first scenario we consider, this is matched by a 4.7 percent increase in the capital share α, shown as the gray dashed line in the figure. In the second scenario, we keep the cost shares of labor and capital fixed at 2/3 and 1/3, implying to a rise in the profit share from 13.5 in 1980 to 20.5 in 2016, or, equivalently, a rise in the markup rate from 15.6 to 25.7 percent (shown as the red dashed line). The resulting capital share decline from 28.8 percent to 26.5 percent is shown as the red solid line.



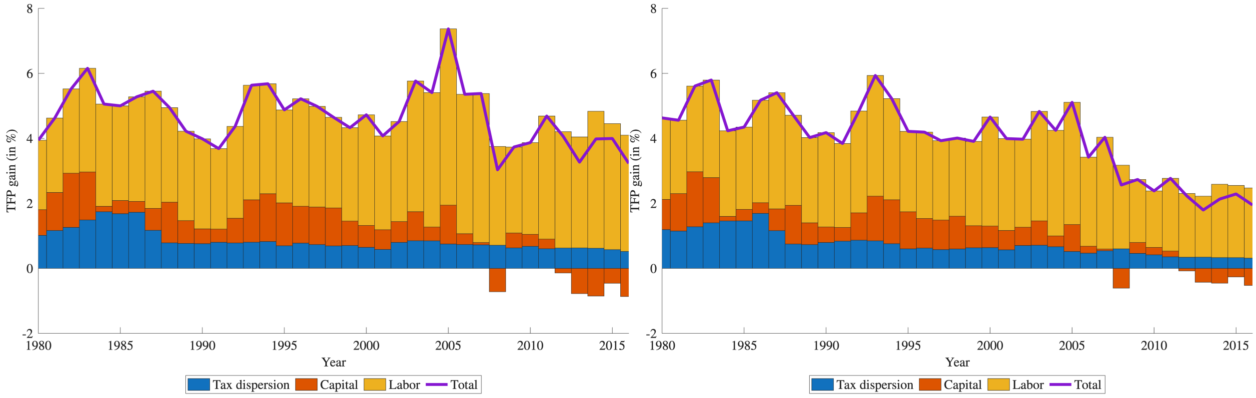
(a) Parameters

(b) Misallocation

Figure A1: Macro trend scenarios

Note.– Panel (a) shows the labor share of income (β) from the BEA, and the associated changes in the shares of capital (α) and profits ($1-\gamma$) under two alternative scenarios. Panel (b) shows the TFP gains from tax rate equalization in each scenario. Parameters are constant in the baseline scenario.

From (7), changes in the labor share of income have a ceteris paribus ambiguous effect on estimated TFP gains. Note that the labor share of income β acts as a weight in equation (7) when considering distortions that are related to the allocation of labor versus capital. A lower β shifts weight away from the correlation between tax distortions and capital distortions. Because the effect of tax heterogeneity on allocative efficiency is generally ambiguous in a distorted economy - it depends on the correlations of tax rates with other capital and labor distortions - the net effect of a change in the labor share is ambiguous as well. For instance, if labor productivity is generally



(a) Rising Capital Share (b) Rising Markups

Figure A2: Sources of TFP gains under alternative macro scenarios

high and capital productivity is generally low among firms that face higher tax rates, then a lower labor share should be associated with smaller TFP losses from tax heterogeneity.

We first consider the scenario of a higher capital share. A higher value of α raises the TFP gains, ceteris paribus. This effect is unambiguous from (7). Intuitively, the larger the importance of the distorted factor in production is, the larger are the losses from tax distortions. However, the dashed gray line in panel (b) of Figure A1 shows that the TFP gains from this scenario yield similar magnitudes as our baseline, indicating that the role of the capital share is quantitatively small.

We now consider the second scenario, under which markups increase, which is equivalent to a lower value of γ. Ceteris paribus, lower values of γ are associated with lower TFP losses from tax heterogeneity as indicated by equation (7). Therefore rising markups would tend to reduce the TFP gains over time. Intuitively, this is because lower values of γ bring the economy further away from a linear technology (or, equivalently, from perfect competition), where the best firm can absorb all the resources without facing diminishing returns to scale. The lack of that possibility lowers the total gains to reallocating inputs more efficiently.

Looking at the red dashed line in panel (b) of Figure A1 once again shows similar TFP gains under this scenario relative to the baseline. Our results point to an approximately 1 percentage point lower potential TFP gain from eliminating tax heterogeneity in recent years.

Decomposition Figure A2 shows the decomposition of TFP gains under the alternative macro scenarios where the labor share declines. The left panel attributes that decline to a higher capital share of income and the right panel to a rise in markups. The relative magnitudes of the interaction between tax wedges and capital or labor distortions are broadly similar across scenarios. Under rising markups, overall TFP gains from eliminating tax heterogeneity decline by more, especially toward the end of the sample. This is attributable to a lower value of γ, which reduces the TFP losses associated with each component (see equation (7)), although the decline in the labor component is the most apparent.

C.4 Sectoral decomposition of TFP gains

This appendix describes the construction of the sector data used in the industry-level TFP loss calculation, how we compute TFP gains at the sector level, and how we aggregate sector-level results to aggregate losses. Those results were reported in section 5.4 of the main paper.

Industry data. We obtain labor and capital shares at the 2-digit NAICS level from the Bureau of Labor Statistics' multifactor productivity data. We compute value added as output minus energy, materials, and services, and define the capital and labor shares as capital costs and labor compensation divided by value added. For every sector, we use a single value for the capital and labor share, which is obtained as the value in the year 2000, roughly the mid-point of our sample period.

In addition, for every sector we compute employment-, capital-, and value-added weights, denoted as λ_N^s, λ_K^s, and λ_Y^s. We use the BLS Current Employment Statistics (CES) data to construct industry-level employment weights. Each 2-digit NAICS industry's employment is measured as aggregate weekly hours of all employees. To obtain capital shares, we use the Bureau of Economic Analysis (BEA) Fixed Assets Accounts private nonresidential fixed assets by industry table. Finally, we compute each industry's share of aggregate value added from these BLS multifactor productivity data. The values of λ_N^s, λ_K^s, and λ_Y^s are taken to be the average across all years in our sample.

Industry-specific TFP gains. To compute industry-specific TFP gains, for each 2-digit sector s, we use the capital and labor shares α^s and β^s from the BLS data and combine them with the common parameter γ to define sector-specific production parameters as $\gamma\alpha^s$ and $\gamma\beta^s$.

For every sector, tax measure, and year, we then compute the TFP gains from eliminating heterogeneity in tax rates, while applying the same firm-level semi-elasticities and tax-wedge distributions as in the baseline. This delivers a sector-level TFP gain from eliminating tax dispersion, g_z^{sjt}, where $j \in \{\text{GM, BCG}\}$ denotes the tax measure.

Finally, we aggregate across sectors using value-added weights λ_Y^s to obtain total productivity gain from reallocating inputs *within* sectors (following tax homogenization in each sector).

$$g_{jt}^{\text{within}} = \sum_s \lambda_Y^s \, g_z^{sjt},$$

Allowing for input reallocation *between* sectors would alter the value-added weights λ_Y^s, as sectors with larger productivity gains, g_z^{sjt}, attract inputs from other sectors. To compute the resulting input shares after reallocation between sectors, note that sectoral production function is of Cobb-Douglas form with respective input elasticities α^s and β^s. Therefore, sector-level quantities are:

$$N^s = Z^{s\,\frac{1}{1-\gamma}} \left(\frac{\beta^s}{\omega_N^s}\right)^{\frac{1-\alpha^s}{1-\gamma}} \left(\frac{\alpha^s}{\Omega_K^s}\right)^{\frac{\alpha^s}{1-\gamma}} \tag{A31}$$

$$K^s = Z^{s\,\frac{1}{1-\gamma}} \left(\frac{\beta^s}{\omega_N^s}\right)^{\frac{\beta^s}{1-\gamma}} \left(\frac{\alpha^s}{\Omega_K^s}\right)^{\frac{1-\beta^s}{1-\gamma}}, \tag{A32}$$

where Ω_K^s and ω_N^s now denote effective sector-level distortions to capital and labor. These equations show that each sector's capital and labor demand grow by $(g_z^{sjt})^{\frac{1}{1-\gamma}}$ in response to improvements in TFP. Using λ'^s_K and λ'^s_N to denote input shares after reallocation, we have:

$$\lambda'^{s}_{K} = K'^{s}/\sum_{s} K'^{s} = \frac{(g_z^s)^{\frac{1}{1-\gamma}} K^s}{\sum_s K^s (g_z^s)^{\frac{1}{1-\gamma}}} = \lambda_K^s \frac{(g_z^s)^{\frac{1}{1-\gamma}}}{\sum_s \frac{K^s}{K} (g_z^s)^{\frac{1}{1-\gamma}}}$$
$$\lambda'^{s}_{N} = N'^{s}/\sum_{s} N'^{s} = \frac{(g_z^s)^{\frac{1}{1-\gamma}} N^s}{\sum_s N^s (g_z^s)^{\frac{1}{1-\gamma}}} = \lambda_N^s \frac{(g_z^s)^{\frac{1}{1-\gamma}}}{\sum_s \frac{N^s}{N} (g_z^s)^{\frac{1}{1-\gamma}}},$$

where $\lambda_K^s = K^s/K$ and $\lambda_N^s = N^s/N$ are a sector's shares of total capital and labor in the distorted economy. If a sector's $(g_z^s)^{\frac{1}{1-\gamma}}$ is above its capital-weighted (labor-weighted) average across all sectors, then that sector obtains a larger share of the capital (labor) stock, i.e. $\lambda'^{s}_{K} > \lambda^s_K$.

Reallocating existing capital and labor supply according to these input shares, $\hat{K}^s = \lambda'^{s}_{K} K$ and $\hat{N}^s = \lambda'^{s}_{N} N$, gives the sectoral output growth after inter-sectoral reallocation of existing inputs as:

$$\frac{\hat{Y}_{sjt}}{Y^s} = \hat{g}^y_{sjt} = g_z^{sjt} \cdot \underbrace{\frac{(g_z^{sjt})^{\frac{\gamma}{1-\gamma}}}{\left(\sum_s \lambda_K^s (g_z^{sjt})^{\frac{1}{1-\gamma}}\right)^{\alpha^s} \left(\sum_s \lambda_N^s (g_z^{sjt})^{\frac{1}{1-\gamma}}\right)^{\beta^s}}}_{\text{Between Sector Input Reallocation}}$$

The first g_z^{sjt} term is the within sector TFP gain, or output gain keeping K^s and N^s constant. The second term shows output gains from between-sector allocation and can be more or less than 1. The total output gain is obtained by aggregating sectoral gains:

$$g_{\text{total}}^{jt} = \sum_s \lambda_Y^s \hat{g}_y^{sjt},$$

where λ_Y^s are sectoral output shares in the distorted economy (as observed in the data). Because the the within and between terms are positively related—the higher is g_z^{sjt}, the higher the input gain from inter-sectoral reallocation—the aggregate gain is larger than the sum of sectoral gains from reallocating inputs within each sector alone. Finally, we refer to the difference between total and within components as the between-sector component of aggregate productivity gains: $g_{jt}^{between} = g_{jt}^{total} - g_{jt}^{within}$. In the main text, we report the average gains obtained by the two tax measures $j \in \{\text{GM, BCG}\}$ across our sample period.