Optimal Taxation of Risky Entrepreneurial Capital

Corina Boar
New York University

Matthew Knowles
City, University of London[†]

August 21, 2023

Abstract

We study optimal taxation in a model with endogenous financial frictions, risky investment and occupational choice, where the wealth distribution affects how efficiently capital is used. The planner chooses linear taxes on wealth, capital and labor income to maximize the steady state utility of a newborn agent. Most agents in the model are poor, leading to an equity motive for taxation. We calibrate the model to the US economy and find low positive levels of optimal capital income and wealth taxes. We express optimal tax rates as a closed-form function of the size of tax bases and their elasticities with respect to tax rates, highlighting the forces behind the result. Because financial frictions are endogenous, higher capital income tax rates tighten financial frictions and reduce output. Thus, optimal capital income taxes are lower than in models with exogenous frictions.

Keywords: entrepreneurship, financial frictions, taxation.

JEL classifications: E2, E6, H2

[*]We thank Teresa Steininger for superb research assistance and our discussants Nils Grevenbrock, Miguel-Angel López-García and Owen Zidar for insightful feedback. For many helpful comments, we thank Mark Bils, Yena Park, and seminar participants at the University of Essex, University of Leicester, University of St Andrews, University of Cologne, University of Bonn, SUNY Binghamton, University of Barcelona, University of Kent, University of Manchester Economic Theory Workshop 2018, London Workshop in Quantitative Macroeconomics 2019, the National Tax Association's 113th Annual Conference on Taxation, and the 2022 European Summer Meeting of the Econometric Society. Support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2126/1 - 390838866 is gratefully acknowledged.

[†]Corresponding author: Matthew Knowles: City, UOL, email: matthew.knowles@city.ac.uk

1 Introduction

The vast literature on optimal capital taxation in general equilibrium typically analyzes models in which all capital is the same and the main cost of capital taxation is its negative effect on aggregate saving. However, critics of capital taxation have long expressed concerns that it has harmful effects not only on the total level of investment, but also on its allocation.[1] The allocative effect arises because taxation may affect incentives for entrepreneurs to take risks,[2] as well as their ability to obtain external finance to fund high-return risky projects, as taxation affects the rate of return external investors are likely to obtain.

Motivated by these concerns, we study the optimal taxation of capital income and wealth in a model in which entrepreneurs undertake risky investments and are subject to financial frictions that arise endogenously due to information frictions. We characterize optimal steady state taxes analytically, which allows us to highlight the important forces at play. We then compute optimal taxes in a calibration of the model to the US economy and show that modeling financial frictions endogenously matters significantly for optimal taxation.

We carry our analysis in a perpetual youth model in which newborn households decide whether to become workers or entrepreneurs. Entrepreneurs, who differ in ability, choose how much capital to allocate to a risky and to a risk-free technology. They can borrow in frictional financial markets, where the friction arises endogenously as a result of entrepreneurs' private information about their idiosyncratic shocks. The effect of the financial friction is that entrepreneurs are limited in their ability to borrow and are unable to fully diversify idiosyncratic risk. This discourages them from allocating capital to the risky technology, which consequently has a higher expected return in equilibrium. Taxes on capital income and wealth affect allocative efficiency by affecting *(i)* how entrepreneurs allocate their capital between the risky and risk-free technologies, *(ii)* how capital is allocated across entrepreneurs of different ability levels, and *(iii)* the fraction of agents who become entrepreneurs. Capital income and wealth taxes are not equivalent because agents who invest in the risky technology earn a higher rate of return to their wealth and pay larger capital income taxes, whereas the wealth tax falls equally on all wealthy agents, regardless of the return they earn.

Our modeling approach builds upon the recent literature on the optimal taxation of entrepreneurs who face financial frictions, but is distinct from almost all of this literature in two key respects.[3] First, our endogenous modeling of financial frictions is a significant departure from the existing literature on the optimal taxation of entrepreneurs, which has most frequently assumed that entrepreneurs' ability to obtain external funds is limited by a

[1] For instance, Hayek (1960, chap. 20) argues that the taxation of profits hinders the accumulation of wealth by entrepreneurs who manage "successful new ventures", preventing them from investing further.

[2] See Cullen and Gordon (2007) and Devereux (2009) and the citations therein.

[3] See, for example, Guvenen et al. (2023), Brüggemann (2021), Panousi and Reis (2014)."Almost all", because some of these issues have been considered by Phelan (2019). We compare our work to this below.

collateral constraint which is not directly affected by taxes. Second, our model is analytically tractable and allows optimal steady state taxes to be written as functions of sufficient statistics, which brings new insights into the important forces at play.

We find that our approach to modeling financial frictions endogenously is consequential for optimal taxation. Differently from the typically assumed exogenous collateral constraint, in our setting, the ability of entrepreneurs to obtain external finance depends on their incentives to misreport their privately observed idiosyncratic shocks, and this is directly affected by tax changes. All else equal, a higher capital income tax encourages entrepreneurs to falsely report that their idiosyncratic shocks were bad, since doing so reduces their measured capital income and therefore their tax burden. In response, financial intermediaries reduce the amount that they provide as external finance until entrepreneurs are willing to honestly report these shocks. As such, a higher capital income tax leads to tighter financial frictions and reduces the capital allocated to the high-return risky technology and aggregate output. This channel is absent from models with collateral constraints or other exogenous financial frictions. As such, we find that our level of optimal capital income tax is lower than in otherwise identical models with exogenous financial frictions.

Despite its complexity, our model is analytically tractable. We show that optimal steady state taxes can be written as a closed form function of the size of the tax base for the various taxes, the degree to which each tax is borne by workers and entrepreneurs, and the partial equilibrium elasticities of the tax bases with respect to each tax, in the spirit of the literature on the 'sufficient statistics' approach to optimal taxation (e.g. Piketty and Saez, 2013).

This analytical tractability generates a number of insights into optimal taxation in models with financial frictions. Crucially, despite the role of frictions and the rich heterogeneity in our setting, the optimal taxation problem reduces to a tradeoff between three considerations. First, raising capital income or wealth tax rates mechanically raises tax revenue, especially from richer entrepreneurs, which can be passed to workers as lower labor income taxes. This is desirable to a utilitarian planner because richer entrepreneurs have a lower marginal utility of consumption than other agents. Second, higher tax rates on capital income and wealth reduce the tax base for these taxes by discouraging saving and the allocation of capital to the high-return risky technology, and so, large increases in these tax rates can reduce tax revenue. Third, higher capital income tax rates, in particular, discourage entry into entrepreneurship. If workers pay a higher total tax per capita than entrepreneurs, then encouraging agents to become workers rather than entrepreneurs increases tax revenue, which can then be passed to workers as lower labor income tax rates.

That the elasticies entering the optimal tax formula are partial equilibrium elasticities, which hold factor prices constant, indicates that the endogenous evolution of factor prices in response to tax changes does not, on the margin, affect optimal tax rates. Rather, the endogenous response of agents' behavior to tax changes *only* matters for optimal taxes insofar as it affects the size of the tax base of each tax and, therefore, tax revenue. Importantly, our perpetual youth structure and financial frictions help ensure that the long-run elasticity of capital income and wealth tax bases with respect to these taxes is finite, in contrast to a frictionless representative agent setting where that elasticity is infinite and optimal steady state capital taxes are driven to zero.

We calibrate the model to match values of tax bases, rates of return and features of financial contracts in the US data, at current US tax rates. The calibrated model generates top wealth inequality similar to the data, as well as a large gap between risky and risk-free rates of return. In our baseline calibration, we find that the optimal capital income tax is 3.7%, the optimal wealth tax is 0.2% and the optimal labor income tax is 28%. Implementing these optimal taxes leads to a long-run consumption equivalent welfare gain of 0.2%.

We find that the elasticities that enter our optimal steady state tax formula are not highly sensitive to changes in tax rates, so that a good approximation to these optimal tax rates can be found by simply calculating the values of the elasticities at the initial steady state (i.e. the status quo), and applying the formula. This is convenient for future research, because in principle the elasticities in our optimal tax formula could be estimated empirically, making it possible to draw conclusions about optimal tax rates without needing to commit to a particular model calibration. Relatedly, we show that our optimal tax formula is robust to a number of modeling features, including the exact details of the financial friction. The specifics of the financial friction are, however, consequential for the value of the elasticity of capital income and wealth with respect to taxes that enters the optimal tax formula. As such, our model also suggests that future empirical work to identify these elasticities would be highly informative about the nature of financial frictions.

Our tax formula also helps explain why we find low positive levels of optimal capital income and wealth taxes. For both taxes, the long-run elasticity of tax bases with respect to the tax rate is large in our calibration, primarily because of compounding effects of taxes on wealth accumulation. These effects are somewhat larger for wealth taxes than capital income taxes for two reasons. First, capital income partly reflects the profits that entrepreneurs earn by taking advantage of the different rates of return of risky and risk-free technologies, and this is less responsive to saving. Second, capital income taxes provide more incentives for agents to become workers, and this raises labor income tax revenue. The effect of these considerations is that, if a planner cared only about wage-earners and not wealth-holders, then the planner would want to rely on capital income rather than wealth taxes. However, this is mitigated by the fact that capital income taxes fall more heavily than wealth taxes on poor entrepreneurs, who have a high marginal utility of consumption, and so the planner wishes to use wealth taxes too so as to to reduce the tax burden on these agents.

Related literature. This paper studies optimal taxation in a model in which taxation affects output via its effect on the allocation of capital. In that sense, our paper builds upon Evans (2015), Shourideh (2014), Itskhoki and Moll (2019), Guvenen et al. (2023), Boar and Midrigan (2020), Basseto and Cui (2020) and Brüggemann (2021). We differ from these papers in two respects. First, we allow for micro-founded financial frictions that arise from asymmetric information, thus allowing for changes in taxes to lead to changes in the tightness of financial frictions. Second, we characterize optimal taxes on capital income and wealth as closed form functions of 'sufficient statistics', which not only enables us to shed light on the tradeoff the planner faces, but also to provide a bridge between theory and empirics.

Within this literature, our paper is closest to Guvenen et al. (2023), who also focus on the different effects of capital income and wealth taxation on the allocation of capital. We differ from Guvenen et al. (2023) along three margins. First, we assume an endogenous financial friction and derive analytical results regarding the determinants of optimal taxes. As we argue in the paper, the nature of the financial frictions is an important determinant of the relative merits of wealth and capital income taxes. Second, in finding optimal taxes we allow the planner to simultaneously choose taxes on wealth, capital income and labor income, rather than restricting it to using only wealth or only capital income taxes in conjuction to labor income taxes. Third, we allow for endogenous occupational choice. Abstracting from the effects of taxes on occupational choice leads us to find it optimal to tax wealth and subsidize capital income, just like Guvenen et al. (2023). However, we show that an endogenous entry margin creates additional incentives to tax capital income rather than wealth and the relative merits of wealth taxes depend on how elastic this margin is.

Our paper is also related to the line of work that quantifies the effect of tax changes in models with entrepreneurs. Examples are Cagetti and De Nardi (2009), Kitao (2008), and Rotberg and Steinberg (2019) who study the effect of changing estate, capital income and wealth taxes in related settings. We differ from this literature in several ways. First, our model is analytically tractable which allows us to make the intuition behind the key mechanisms as transparent as possible. Second, our financial friction arises endogenously as a consequence of asymmetric information and is itself affected by changes in taxes, which we show is of importance when considering optimal taxation.

Our paper also relates to studies of optimal taxation in the presence of idiosyncratic investment risk, such as Panousi and Reis (2014), Panousi and Reis (2019) and Phelan (2019). In these papers, unlike our setting, entrepreneurs do not differ in their productivity levels and the allocation of capital does not affect aggregate output. Furthermore, we allow for endogenous effects of tax changes on financial frictions, which mitigate the benefits of capital income taxation. By studying optimal taxation with endogenous financial frictions, our paper also relates to Dávila and Hébert (2023), who consider optimal corporate taxation, where corporations differ in productivity and financial frictions arise due to a limited enforcement constraint. While the financial friction differs, taxes affect the efficiency of capital allocation in a similar way to ours. This work differs substantially from ours in focusing on corporations, rather than entrepreneurs, and not incorporating household heterogeneity.

Lastly, our paper contributes to the literature on optimal capital taxation, which focuses on the effect of capital taxation on aggregate capital accumulation, as in the work of Chamley (1986), Judd (1985), Straub and Werning (2020), Benhabib and Szőke (2021), Chen et al. (2019), among others.[4] Related to our paper, Abo-Zaid (2014), Biljanovska (2019) and Biljanovska and Vardoulakis (2019) explore how the results in this line of work are affected in settings with reduced-form financial frictions while maintaining the assumptions of Chamely and Judd that capital is homogeneous and there is no idiosyncratic risk.

The rest of the paper is organized as follows. Section 2 outlines the assumptions of the model. Section 3 discusses properties of the model equilibrium. Section 4 shows how the steady state is affected by taxes and derives formulae for the optimal tax rates. Section 5 shows the values of optimal taxes in the numerical calibration. Section 6 concludes.

2 Model

In this section we describe our model economy and define an equilibrium.

Environment The economy is populated by a unit mass of households and competitive banks. Households are born identical and with no wealth. At birth each household chooses whether to be an entrepreneur or a worker and retains this occupation for their entire life. Workers supply labor inelastically. Entrepreneurs use capital to produce intermediate goods. Each entrepreneur owns two investment projects: a risky project that produces 'risky' intermediate goods y_E, and a risk-free project that produces 'risk-free' intermediate goods y_F.[5] Entrepreneurs use labor and intermediate goods to produce a final good. The government levies taxes and funds exogenous government spending $\overline{G}$.

Timing Each period is divided into three sub-periods: morning, afternoon and evening. In the morning, entrepreneurs trade capital and divide it between risky and risk-free projects. In the afternoon, they draw idiosyncratic shocks which affect the capital in the risky project. The the projects produce intermediate goods, which entrepreneurs trade. In the evening, they use intermediate goods and labor to produce the final good. Households divide their resources between consumption and saving for the next period. At the end of the period, a fraction γ of households die and new households are born. Capital depreciates at rate δ.

[4] See Chari and Kehoe (1999) for a survey.

[5] The device of having two types of intermediate goods is a simple way to incorporate misallocation of capital into the model and to allow entrepreneurs to choose between more and less risky allocations of capital.

Technology of Entrepreneurs Newborn entrepreneurs draw ability $\theta \in [0,1]$ from a distribution H_θ. At the beginning of each period, an entrepreneur retains the same θ as in the previous period with probability $1-\lambda_\theta$ and draws a new θ from H_θ with probability λ_θ.

After allocating capital between the risky and risk-free projects in the morning, each entrepreneur i draws a idiosyncratic shock $\xi_{i,t}$ from a distribution H_ξ, with mean zero, standard deviation one and full support on $\mathbb{R}$.[6] Therefore, an entrepreneur who allocates $k_{E,i,t}$ to the risky project in the morning has $\tilde{k}_{E,i,t} = q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t})$ units of capital in the project in the afternoon, where q is increasing in $k_{E,i,t}$ and $\xi_{i,t}$. Each unit of capital $\tilde{k}_{E,i,t}$ produces one unit of risky intermediate goods $y_{E,i,t}$. The risk-free project produces an output of $y_{F,i,t} = k_{F,i,t}$ risk-free intermediate goods. We assume that

$$q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t}) = \begin{cases} k_{E,i,t} & \text{if } k_{E,i,t} \leq \underline{k}_E \\ k_{E,i,t} + (1-\underline{\epsilon})\left(\exp\left(\frac{\varphi \xi_{i,t}}{\sqrt{\theta_{i,t}}} - \frac{\varphi^2}{2\theta_{i,t}}\right) - 1\right)(k_{E,i,t} - \underline{k}_E) & \text{if } k_{E,i,t} > \underline{k}_E \end{cases}$$

where $\underline{\epsilon} \in (0,1)$, $\varphi > 0$ and $\underline{k}_E > 0$. This implies that an entrepreneur can put $k_{E,i,t} \leq \underline{k}_E$ into the risky project without facing any risk. If $k_{E,i,t} > \underline{k}_E$ the project becomes risky: the mean of $\tilde{k}_{E,i,t}$ is equal to $k_{E,i,t}$, and its variance is decreasing in the entrepreneur's ability and is increasing and convex in the level of investment in this technology. The former property implies that having higher ability reduces the risk that entrepreneurs face for a given amount of capital in the risky project. The latter is akin to decreasing returns to scale in production and guarantees a positive measure of entrepreneurs in the steady state.[7] The lowest realization of $\tilde{k}_{E,i,t}$ is $(1-\underline{\epsilon})k_{E,i,t} + \underline{\epsilon}\underline{k}_E$.

Our specific assumption on the functional form of q, although not critical for our derivation of the optimal tax formula, has two convenient properties. First, it implies that optimal contracts for obtaining external funds are identical to equity and debt contracts, the most common financial contracts in the data. This helps with calibrating the agency frictions against the data. Second, it enables us to study the effects of taxes on the allocation of capital in a tractable way, while allowing for heterogeneity in entrepreneurial ability.

Technology of Final Good Production Entrepreneurs trade risky intermediate goods at price $r_{E,t}$ and risk-free intermediate goods at price $r_{F,t}$. Each entrepreneur i hires $n_{i,t}$ workers at wage rate w_t and uses $y_{E,i,t}$ and $y_{F,i,t}$ units of intermediate goods to produce $y_{i,t}$ final goods with technology $y_{i,t} = f(y_{E,i,t}, y_{F,i,t}, n_{i,t})$, where f is concave and strictly increasing in all arguments, exhibits constant returns to scale and satisfies the Inada conditions.

[6] These restrictions on the first two moments of H_ξ and the upper bound on θ are normalizations.

[7] Convexity of entrepreneurial risk with respect to risky capital invested is more tractable in our setting than having decreasing returns to scale in production, but has similar implications: *(i)* high ability entrepreneurs cannot produce too much and thus generate too high profits and *(ii)* there are economic rents that increase the profits of poor entrepreneurs and make this occupation more attractive.

Hiding Capital Entrepreneurs can hide capital $k_{H,i,t}$ in their risky project after observing the shock $\xi_{i,t}$ and convert it into $\phi k_{H,i,t}$ units of consumption, where $\phi \in (0,1)$.[8] We will show that, when taxes are set optimally, entrepreneurs will not choose to hide any units of capital. However, the ability to hide capital creates *endogenous* frictions in financial markets.

Preferences Households born in period t maximize expected lifetime utility given by $\sum_{j=1}^{\infty} (1-\rho)^{j-1} (1-\gamma)^{j-1} u_{i,t+j}$, where $u_{i,t+j}$ is household i's period utility in period $t+j$. For a worker, $u_{i,t+j} = \log(c^N_{i,t+j})$. For an entrepreneur, $u_{i,t+j} = \log(c_{i,t+j}) + z_i$, where $z_i \in \mathbb{R}$ captures the household-specific non-pecuniary benefits of being an entrepreneur (see Hurst and Pugsley, 2011). At birth, households draw z_i from the distribution H_z. We impose that $\varphi^2 > \lambda_\theta + \rho + 2\gamma$, which holds easily in our calibration and guarantees that the allocation of capital to risky projects is proportional to entrepreneurial ability, simplifying aggregation.

Occupational Choice Newborn households choose their occupation to maximize expected lifetime utility. There exists a cutoff $z_t^\star$ such that only individuals with $z_i \geq z_t^\star$ will choose to be entrepreneurs. The cutoff $z_t^\star$ satisfies

$$\sum_{j=1}^{\infty} (1-\rho)^{j-1} (1-\gamma)^{j-1} \log(c^N_{i,t+j}) = \mathbb{E}_t \left[\sum_{j=1}^{\infty} (1-\rho)^{j-1} (1-\gamma)^{j-1} \left(\log(c_{i,t+j}) + z_t^\star \right) \right],$$

where the expectation is with respect to the future realizations of $\theta_{i,t}$ and $\xi_{i,t}$.

Government The government levies a labor income tax $\tau_{N,t}$, a capital income tax $\tau_{K,t}$ and a wealth tax $\tau_{W,t}$, and has to finance exogenous expenditure $\overline{G}$. The government's budget constraint each period is

$$\overline{G} \;=\; \tau_{N,t} w_t N_t + \tau_{K,t}(\Pi_t - \delta K_t) + \tau_{W,t} K_t, \tag{1}$$

where N_t is the measure of workers, K_t is the aggregate capital stock at the start of the period and $\Pi_t - \delta K_t$ are the total *reported* profits of entrepreneurs net of capital depreciation.

Financial Markets Entrepreneurs can fund capital purchases by entering one-period state-contingent contracts with risk neutral banks that live for one period. An entrepreneur who borrows $b_{i,t} > 0$ in the morning returns $\hat{b}_{i,t}$ in the evening. A bank will only lend to entrepreneurs if the expected return on the loan exceeds the market risk-free rate $R_{F,t}$

$$\mathbb{E}_\xi \hat{b}_{i,t} \geq R_{F,t} b_{i,t},$$

where the expectation with respect to the realizations of ξ. In equilibrium, this condition holds with equality and banks make zero profits. Workers can also borrow at rate $R_{F,t}$.

[8] As we subsequently discuss, the realization of $\xi_{i,t}$ is private information to the entrepreneurs.

Annuities At the end of the period households trade financial annuities to insure against the risk of death. A household can exchange a unit of the final good for the promise of receiving $\frac{1}{1-\gamma}$ units at the start of the next period if still alive. Entrepreneurs place all their capital in a common fund at the end of the period, exchanging it for annuities.[9]

Budget Constraints The budget constraint of a worker with wealth $a_{i,t}^N$ is[10]

$$c_{i,t}^N + (1-\gamma)a_{i,t+1}^N = w_t(1-\tau_{N,t}) + R_{F,t}a_{i,t}^N.$$

In the morning the budget constraint of the entrepreneur is

$$k_{E,i,t} + k_{F,i,t} = k_{i,t} = a_{i,t} + b_{i,t}.$$

After receiving the $\xi_{i,t}$ shock, the entrepreneur chooses how many units of capital $k_{H,i,t}$ in the risky project to hide and transform into consumption $c_{H,i,t}$. In the evening, she chooses consumption $c_{i,t}$ and annuities $(1-\gamma)a_{i,t+1}$, repays the bank $\hat{b}_{i,t}$ and pays taxes. Consequently, in the evening the entrepreneur's budget constraint is

$$c_{i,t} - c_{H,i,t} + (1-\gamma)a_{i,t+1} + \hat{b}_{i,t} = \pi_{i,t} - T_{i,t} + (1-\delta)k_{i,t},$$

where $\pi_{i,t}$ denotes the entrepreneur's reported period profits given by

$$\begin{aligned}\pi_{i,t} &= \underbrace{(r_{E,t}y_{E,i,t} + r_{F,t}y_{F,i,t})}_{\text{profit from intermediate goods}} + \underbrace{(y_{i,t} - w_t n_{i,t} - r_{E,t}y_{E,i,t}^d - r_{F,t}y_{F,i,t}^d)}_{\text{profit from final good}} \\ &\quad + \underbrace{(1-\delta)(\tilde{k}_{E,i,t} - k_{H,i,t} - k_{E,i,t})}_{\text{reported capital gain}},\end{aligned}$$

and where $c_{H,i,t} = \phi k_{H,i,t}$, $y_{E,i,t} = \tilde{k}_{E,i,t}$ and $y_{F,i,t} = k_{F,i,t}$. We assume that hidden capital is not reported as profits for tax purposes since it is hidden from outside agents, including the government. Given the constant returns to scale technology, the profits from the final good are zero in equilibrium. Finally, the tax payment $T_{i,t}$ is equal to $\tau_{K,t}\pi_{i,t} - \tau_{K,t}\delta k_{i,t} + \tau_{W,t}k_{i,t}$.

Agency Friction An entrepreneur's realization of ξ, the capital that she hides and the consumption she obtains from doing so are all private information.[11] Without loss of generality, we restrict attention to incentive compatible contracts where the entrepreneur truthfully reports her ξ and pays the bank the promised amount $\hat{b}$. This gives rise to the

[9] Even if allowed to hold capital, they would prefer the common fund, since it insures against death risk.

[10] Workers do not pay capital income or wealth taxes. For analytical convenience and without loss of generality, we assume such taxes are levied on physical assets and profits only, not on financial assets.

[11] When $\phi = 0$ there is no informational friction, since there is no incentive to hide capital.

following incentive compatibility constraint

$$\underbrace{(1-\tau_K)\,(r_E+1-\delta)\frac{\partial\tilde{k}_E}{\partial\xi}}_{\text{marginal cost of under-reporting }\xi} \geq \underbrace{\phi\frac{\partial\tilde{k}_E}{\partial\xi}+\frac{\partial\hat{b}}{\partial\xi}}_{\text{marginal benefit of under-reporting }\xi}. \tag{2}$$

By under-reporting ξ by a small amount $\mathrm{d}\xi$, the entrepreneur can hide $\frac{\partial\tilde{k}_E}{\partial\xi}\mathrm{d}\xi$ units of capital. This means that she produces this many fewer units of risky intermediate goods and loses the after-tax return from selling them. At the same time, she converts the hidden units of capital into $\phi\frac{\partial\tilde{k}_E}{\partial\xi}\mathrm{d}\xi$ units of consumption and also repays less to the bank.

Worker's Problem Letting X denote the aggregate state of the economy, the worker chooses consumption $c^N(a^N,X)$ and annuities $a^{N\prime}(a^N,X)$ that solve the Bellman equation

$$V^N(a^N,X)=\max_{c^N,a^{N\prime}}\ \log\left(c^N\right)+(1-\rho)\,(1-\gamma)\,V^N\left(a^{N\prime},X'\right),$$

subject to the worker's budget constraint.

Entrepreneur's Problem Letting $V(a,\theta,X)$ denote the value of an entrepreneur with wealth a and ability θ, the expected lifetime utility of a newborn entrepreneur i is

$$\int_0^1 V(0,\theta,X)\mathrm{d}H_\theta(\theta)+\frac{z_i}{1-(1-\rho)(1-\gamma)}.$$

Since the second term is a constant, the entrepreneur solves the Bellman equation

$$V(a,\theta,X)=\sup\int_\xi\Big(\log(c(a,\theta,\xi,X))+(1-\rho)(1-\gamma)\mathbb{E}\big[V(a'(a,\theta,\xi,X),\theta',X')\big|\theta\big]\Big)\mathrm{d}H_\xi(\xi),$$

subject to the morning and evening budget constraints, the production functions for c_H, y_E, y_F, and y, the incentive compatibility constraint and the banks' break-even condition.

Equilibrium Given a sequence of tax rates $\{\tau_{W,t},\tau_{K,t},\tau_{N,t}\}_{t=0}^\infty$, an equilibrium is a sequence of prices $\{R_{F,t},r_{E,t},r_{F,t},w_t\}_{t=0}^\infty$ and decision rules of entrepreneurs and workers such that newborn agents choose the occupation that maximizes expected lifetime utility, workers' and entrepreneurs' decision rules solve their respective optimization problems, the government's budget is balanced every period, and the asset, labor intermediate goods and final goods markets clear every period.

3 Properties of the Model Equilibrium

In this section, we discuss the properties of the model equilibrium. We show that the optimal contract between entrepreneurs and banks gives rise to financial frictions that vary endogenously with taxes, and we describe the optimal choices of workers and entrepreneurs.

3.1 Optimal Contract

We first show that the optimal financial contract between the entrepreneur and the bank has an easily interpretable form as an equity and debt contract. The share of equity that the entrepreneur must retain in her project varies *endogenously* with taxes, a feature we later show to matter in determining optimal taxes.

To that end, we note that the entrepreneur's problem can be split into a within-period choice of maximizing end-of-period resources by allocating capital across projects, borrowing and hiding capital, and a between period choice of dividing these resources between c and a'. We formally state and solve these problems, as well as the problem of workers, in Appendix A. The solution to the between period problem yields the entrepreneur consuming a constant share $1-(1-\rho)(1-\gamma)$ of the end-of-period resources and saving the rest. The within-period problem takes the form of a standard portfolio choice problem, with a trade-off between risk and return: choosing a higher investment in the risky project increases the variance of end-of-period resources, but also their expected value.

Letting ω denote the end-of-period resources of the entrepreneur, the incentive compatibility constraint (2) can be rewritten as

$$\frac{\partial \omega}{\partial \xi} \geq \phi \frac{\partial \tilde{k}_E}{\partial \xi}. \tag{3}$$

By under-reporting ξ by a small amount $\mathrm{d}\xi$, the entrepreneur can hide $\frac{\partial \tilde{k}_E}{\partial \xi}\mathrm{d}\xi$ units of capital and convert them into $\phi\frac{\partial \tilde{k}_E}{\partial \xi}\mathrm{d}\xi$ units of consumption. This reduces end-of-period resources by $\frac{\partial \omega}{\partial \xi}\mathrm{d}\xi$, representing the cost of forgoing the after-tax return from the intermediate goods that could have produced, net of the benefit of having to repay less to the bank.

Integrating equation (3) with respect to ξ, it follows that there exists a function $\underline{\omega}$ that depends on the entrepreneurs wealth, ability θ and prices such that

$$\omega \equiv \underline{\omega} + \phi(\tilde{k}_E - \underline{k}_E). \tag{4}$$

In the absence of agency frictions, the risk-averse entrepreneur and the risk-neutral bank would prefer a contract in which the bank takes all the risk and the entrepreneur's ω is independent of ξ. The agency friction prevents this, leading the entrepreneur to face the level of risk implied by equation (4). This uniquely pins down the value of $\hat{b}$ in each state of the world. The resulting contract between the entrepreneur and bank takes an easily interpretable form as an equity and debt contract, as discussed in the following Lemma.

Lemma 1. *The equilibrium financial contract is one in which the entrepreneur takes a loan less than or equal to fraction R_F^{-1} of the end of period value of her risky project under the worst possible realization of ξ, and sells fraction $1 - \frac{\phi}{(1-\tau_K)(r_E+1-\delta)}$ of the remaining value of her investment projects as equity, retaining the fraction $\frac{\phi}{(1-\tau_K)(r_E+1-\delta)}$ of the equity herself.*

Proof. See Appendix A.3 □

The reason the entrepreneur cannot sell all the equity in her projects is that she needs to have a large enough 'skin in the game' to prevent her from hiding capital. The share of the project she must retain varies endogenously with taxes: a higher capital income tax reduces the fraction of equity the entrepreneur is able to sell, thus tightening the financial frictions.

The Inada conditions on production imply that entrepreneurs must put some capital into the risk-free technology, and produce some risky intermediate goods. No-arbitrage implies

$$R_F = 1 + (1-\tau_K)(r_F - \delta) - \tau_W, \tag{5}$$

$$\phi \leq (1-\tau_K)(r_E + 1 - \delta), \text{ with equality if } k_H > 0, \tag{6}$$

$$0 < (r_E - r_F)(1-\tau_K). \tag{7}$$

Equation (5) states that the risk-free return to lending to a bank equals the return to putting capital in the risk-free technology. Equation (6) states that hiding capital cannot be more lucrative than selling risky intermediate goods. Equation (7) implies that the risky technology has a higher return than the risk-free technology, to compensate for risk. In equilibrium, each entrepreneur chooses $k_E \geq \underline{k}_E$, since borrowing and investing up to $\underline{k}_E$ is possible without risk, and equation (7) implies that this yields a positive return.

3.2 Optimal Decisions of Workers and Entrepreneurs

We next describe the optimal choices of workers and entrepreneurs as functions of prices and taxes. By aggregating these choices we can analyze how the steady state of the economy responds to changes in taxes. To do this, as well as to ultimately find the optimal taxes, it is analytically more convenient to work with the continuous time version of our economy, which we solve formally in Appendix B. In brief, the continuous time version of the model is obtained by assuming each period is of time length Δ and taking the limit as Δ approaches zero. Here we directly discuss the resultant optimal choices.

In equilibrium, all agents consume a constant fraction of their resources and save the rest. Specifically, letting P^N denote the expected present discounted value of lifetime labor income, net of labor income taxes, the optimal choice of a worker is given by

$$c^N = (\rho + \gamma)P^N \qquad \text{and} \qquad \mathrm{d}P^N = \left\{ \left[\tilde{R}_F + \gamma \right] P^N - c^N \right\} \mathrm{d}t,$$

where $\tilde{R}_F = R_F - 1$ denotes the net risk-free rate of return. Similarly, letting P denote the expected present discounted value of the lifetime income of an entrepreneur who takes no risk,[12] the consumption-saving choice of an entrepreneur with ability θ is given by

$$\begin{aligned} c &= (\rho + \gamma)P \\ \mathrm{d}P &= \left[\left(\tilde{R}_F + \gamma \right) P + (k_E - \underline{k}_E)(r_E - r_F)(1 - \tau_K) - c \right] \mathrm{d}t + \frac{(k_E - \underline{k}_E)\phi(1 - \underline{\epsilon})\varphi}{\sqrt{\theta}} \mathrm{d}W, \end{aligned}$$

where $\mathrm{d}W$ is the difference of a standard Brownian motion. In Appendix B we also show that if $r_E > r_F$, then entrepreneurs do not hide capital and so $k_H = 0$. This is because the return to hiding capital is always lower than the return to selling intermediate goods, i.e. the inequality (6) is strict.

We also characterize how the allocation of an entrepreneur's capital between the risky and risk-free projects varies with taxes and the severity of financial frictions. The amount of capital the an entrepreneur of type θ invests in the risky technology is

$$k_E = \underline{k}_E + P\hat{k}_E(\theta),$$

where

$$\hat{k}_E(\theta) \equiv \frac{1}{\phi(1 - \underline{\epsilon})} \times \min \left[\frac{(r_E - r_F)(1 - \tau_K)\theta}{\phi(1 - \underline{\epsilon})\varphi^2}; 1 \right].$$

This shows that, all else equal, richer entrepreneurs invest more in risky projects. Furthermore, an entrepreneur's investment in risky projects is closely tied to the Sharpe ratio they face for risky projects, which is equal to $(r_E - r_F)(1 - \tau_K) / \left(\frac{\phi^2 (1-\underline{\epsilon})^2 \varphi^2}{\theta} \right)$. The Sharpe ratio is higher when *(i)* the after-tax return to risky projects is higher, *(ii)* the after-tax return to risk-free projects is lower or *(iii)* the agency friction is less severe (i.e. lower ϕ).

An implication of this is that capital income taxes reduce incentives to invest in risky projects. This is a consequence of our endogenous financial friction, and arises because capital income taxes reduce the post-tax excess return for risky projects $(r_E - r_F)(1 - \tau_K)$, but not the risk that entrepreneurs face for investing in these projects. The latter might seem surprising because capital income taxes, by taxing capital gains, do reduce the variance

[12] Such an entrepreneur puts exactly $\underline{k}_E$ units of capital into the risky technology each period, no capital into the risk-free technology, and lends her remaining wealth $a - \underline{k}_E$ to banks at the risk-free rate R_F.

of entrepreneurial profits . However, the level of risk entrepreneurs ultimately bear is that required to persuade them to honestly report their ξ shocks, as given by equation (4). Conditional on their risky investment, this is not directly affected by taxes. Therefore, insofar as the tax system insures entrepreneurs against risk, this tightens incentive compatibility constraints and crowds out the insurance against risk provided by financial intermediaries. As such, the fraction of equity entrepreneurs can sell externally in Lemma 1 is decreasing in τ_K, and so a higher τ_K reduces their ability to insure against risk through this channel. Then, since a higher τ_K reduces the excess return from risky projects but does not reduce their risk to the entrepreneur, it shifts investment away from risky projects. This channel is a consequence of our endogenous financial friction, and, as we discuss in Section 5.3.2 does not arise in otherwise identical models with various exogenous financial frictions.

Lastly, as discussed above, all newborn agents who draw non-pecuniary benefits z_i higher than a cutoff $z^\star$ choose to be entrepreneurs. The measure of workers N evolves according to

$$\frac{\partial N}{\partial t} = -\gamma N + \gamma \text{Prob}(z_i < z^\star) = -\gamma N + \gamma H_z(z^\star).$$

The first term captures outflows due to stochastic death, while the second captures inflows of newly born agents who draw non-pecuniary shocks below the cutoff. Because taxes affect the lifetime value of being in a given occupation, the cutoff $z^\star$ responds endogenously to taxes, so taxes also affect occupational choice.

4 Optimal Taxes

In this section, we characterize the effect of taxes on the steady state of the economy and then solve for the taxes that maximize the steady-state welfare of a newborn agent. We obtain a formula for optimal taxes, which we show is robust to many features of our model.

4.1 The Effect of Taxes on Capital Allocation

Before calculating optimal taxes, we first discuss the tradeoff that the planner faces in designing optimal policy. We focus the discussion on the effect that capital income and wealth taxes have on the accumulation of capital and its allocation across risky and risk-free investments, as these play a key role in determining the value of optimal taxes.

We let K and K_E denote the steady state values of the aggregate capital stock and the capital in the risky technology, obtained by integrating over the individual choices of entrepreneurs with respect to the distribution of wealth across entrepreneurs.[13] To study

[13] These are fomally characterized in Appendix B.4, together with the other quantities and prices that constitute the steady-state of the economy.

how these respond to changes in τ_K and τ_W, we compute their partial equilibrium elasticity with respect to taxes, holding constant the pre-tax prices r_E, r_F and w, but allowing for endogenous changes in the distribution of wealth. While holding r_F fixed, we incorporate that R_F is affected by tax changes according to the no-arbitrage condition (5). To emphasize the effect that financial frictions have on the determination of optimal taxes (due to their effects on the accumulation and allocation of capital), we assume that the planner chooses τ_K and τ_W, and adjusts the labor income tax τ_N to balance its budget. We focus on these partial equilibrium elasticities because, as we show below, optimal taxes can be written as functions of these elasticities, ignoring general equilibrium effects through prices, as in Diamond and Mirrlees (1971) and Piketty and Saez (2013).

We define the partial equilibrium elasticities of an aggregate variable X with respect to the tax rates τ_K and τ_W as[14]

$$e^X_{\tau_K} \equiv \frac{(1-\tau_K)}{X}\frac{\partial X}{\partial \tau_K} \qquad \text{and} \qquad e^X_{\tau_W} \equiv \frac{1}{X}\frac{\partial X}{\partial \tau_W}.$$

By considering small perturbations in post-tax prices around the steady-state, we show in Appendix B.5 that the partial equilibrium elasticity of steady state capital in the risky technology K_E with respect to the tax rates τ_j, $j \in \{K, W\}$ is

$$e^{K_E}_{\tau_j} \;=\; \left(1 - \frac{\underline{k}_E\,(1-N)}{K_E}\right) M_{K_E}(e^{\hat{k}_E}_{\tau_j} + e^{\mathbb{P}}_{\tau_j}) - \frac{N e^N_{\tau_j}}{1-N}.$$

This elasticity provides intuiton for how capital income and wealth taxes affect risky investment. In particular, a change in taxes has three effects on K_E, captured by $e^{\hat{k}_E}_{\tau_j}$, $e^{\mathbb{P}}_{\tau_j}$ and $e^N_{\tau_j}$. First, the term $e^{\hat{k}_E}_{\tau_j}$ reflects that a tax change affects entrepreneurs' incentives to invest capital into the risky technology by changing their choice of $\hat{k}_E$ and the distribution of wealth across entrepreneurs. Second, the term $e^{\mathbb{P}}_{\tau_j}$ captures the effect on the lifetime resources of entrepreneurs and is negative. An increase in τ_K or τ_W reduces these resources, conditional on K_E, through directly reducing entrepreneurs' income and encouraging consumption. These two terms are multiplied by $\left(1 - \frac{\underline{k}_E(1-N)}{K_E}\right) M_{K_E}$. The term M_{K_E} captures the multiplier effect that arises because a higher K_E increases entrepreneurs' wealth, thus raising K_E further. The term $1 - \frac{\underline{k}_E(1-N)}{K_E}$ captures that changes in entrepreneurs' choices and wealth only affect the part of K_E over and beyond the $\underline{k}_E$. Thus, when $\frac{\underline{k}_E(1-N)}{K_E}$ is close to 1 all entrepreneurs put roughly $\underline{k}_E$ capital into the risky technology so K_E is inelastic in response to taxation. Lastly, the term $e^N_{\tau_j}$ captures the fact that an increase in τ_K or τ_W tends to shift households to becoming workers rather than entrepreneurs, which reduces K_E.

The elasticities $e^{\hat{k}_E}_{\tau_K}$ and $e^{\hat{k}_E}_{\tau_W}$ are particularly relevant for how taxes affect the allocation of

[14]In the case of a wealth tax, this is a semi-elasticity.

capital, since they determine how entrepreneurs allocate their capital to the risky technology. In Appendix B.5 we show that they are approximately -1 and 0 respectively, provided λ_θ is sufficiently large. This indicates that capital income taxes have additional important effects on the allocation of capital, which wealth taxes do not. The most important of these effects is that higher taxes on capital income encourage entrepreneurs to reduce their risky investment by reducing the excess return to this investment but not the risk to the entrepreneur in the equilibrium financial contract, as discussed in Section 3.2. As we discuss in Section 5.3.2, the elasticity $e^{\hat{k}_E}_{\tau_K}$ is strongly affected by the endogenous financial frictions assumed in the model, and can be quite different when financial frictions are modeled exogenously.

Similarly, we obtain a formula for the partial equilibrium elasticity of steady state capital stock K with respect to the tax rates τ_j, $e^{K}_{\tau_j}$, $j \in \{K, W\}$, as shown in Appendix B.5. We omit the formula here, but note that a change in τ_K or τ_W also has three effects on K. First, the classic argument by which taxes affect the post-tax return to saving applies, and an increase in taxes favors consumption over saving. Second, as discussed above, taxes affect K_E. In turn, this increases saving, and thus the capital stock, by increasing entrepreneurs' income. Third, by affecting the value of being an entrepreneur and therefore occupational choice, taxes increase saving by workers and reduce consumption by entrepreneurs.

In Appendix B.5, we also characterize the partial equilibrium elasticity of output Y and the mass of workers N with respect to taxes, $e^{Y}_{\tau_j}$ and $e^{N}_{\tau_j}$, $j \in \{K, W\}$. Like in the neoclassical growth model, taxes affect output by affecting capital accumulation and the mass of workers. In addition, in our environment with financial frictions, taxes also affect output by reallocation capital between the risky and the risk-free technology. An increase in output due to this reallocation effect ultimately represents an increase in aggregate productivity, since it corresponds to an increase in output with no increase in the factors of production.

4.2 The Effect of Taxes on Welfare

To calculate optimal taxes, we use a perturbation approach, which requires that we first characterize the marginal effects of changes in tax rates on welfare and set them equal to zero to recover the optimum. The measure of welfare we consider is the present discounted lifetime utility of a newborn agent in the steady state, denoted by $\mathcal{W}$. To construct this measure, recall that a newborn agent chooses the occupation that maximizes lifetime utility, given the draw of the non-pecuniary benefit of being an entrepreneur. All newborn agents who draw non-pecuniary benefits higher than a cutoff $z^\star$ choose to be entrepreneurs. Therefore, the cutoff $z^\star$ equates the value of being a worker with the expected value of being an entrepreneur

$$V^N\left(P^N, X\right) = \frac{z^\star}{\rho + \gamma} + \mathbb{E}_\theta V\left(P, \theta, X\right),$$

where $V^N\left(P^N, X\right)$ denotes the value of being a worker with lifetime resources P^N. Then, it follows that the expected lifetime utility of a newborn agent is

$$\mathcal{W} = V^N\left(F^N, X\right) + \frac{1}{\rho+\gamma}\int_{z^\star}^{\infty}\left(z - z^\star\right)\mathrm{d}H_z\left(z\right).$$

Taking a first order approximation and using that $N = H_z(z^\star)$, it follows that the effect on welfare of a marginal change in tax rates is[15]

$$\mathrm{d}\mathcal{W} = \mathrm{d}V^N\left(F^N, X\right) - \frac{\mathrm{d}N}{\rho+\gamma}\left(\frac{1 - H_z\left(z^\star\right)}{H_z'\left(z^\star\right)}\right).$$

The first term represents the change in the welfare of workers resulting from the tax reform. This can be shown to be proportional to the change in worker income due to changes in the post-tax prices $w\left(1-\tau_N\right)$ and R_F. The second term represents the fact that, if the tax reform increases the number of workers, then it must be increasing the cutoff $z^\star$, and therefore making entrepreneurs relatively worse off, so the increase in aggregate welfare is less than the increase in worker welfare.

Choosing taxes optimally means that at the optimal tax rates $\mathrm{d}\mathcal{W} = 0$ for a small change in taxes. Similar to Diamond and Mirrlees (1971) and Piketty and Saez (2013), the first order condition for optimal taxes can be formulated as one where $\mathrm{d}\mathcal{W} = 0$ for a small tax change $\mathrm{d}\tau_j$, $j \in \{K, W\}$, holding constant the pre-tax prices r_E, r_F and w, with τ_N determined by budget balance, and R_F determined by the no-arbitrage condition (5). This is because the government has three tax instruments, τ_K, τ_W and τ_N and so can set these to target the values of three post-tax prices (essentially the values of r_E, r_F and w after taxes are deducted).[16] Therefore, the optimal policy is one where a small change in post-tax prices leads to no change in welfare on the margin, which is the same as saying that a small change in taxes leads to no change in welfare on the margin, holding fixed pre-tax prices.

4.3 Optimal Tax Formula

We derive a general formula for optimal taxes that depends only on the size of tax bases in the economy and the partial equilibrium elasticities of these tax bases with respect to taxes.

Tax Formula. To derive the formula, let B_{τ_j} denote the tax base for the tax τ_j, so that

$$B_{\tau_N} = wN \qquad \text{and} \qquad B_{\tau_K} = \left(r_E - r_F\right)K_E + \left(r_F - \delta\right)K \qquad \text{and} \qquad B_{\tau_W} = K.$$

[15] We further characterize this in terms of elasticities in Appendix B.6.

[16] The reason that R_F must be considered to evolve according to (5) is that this is a fourth price, while the government only has three tax instruments to target three prices. However, (5) implies that R_F can be directly calculated from taxes and r_F, without needing to consider other general equilibrium effects.

We can then easily calculate elasticities of the tax bases B_{τ_m}, $m \in \{K, W, N\}$, with respect to taxes τ_j

$$e_{\tau_j}^{B_{\tau_m}} = \frac{1}{B_{\tau_m}} \frac{\partial B_{\tau_m}}{\partial \tau_j},$$

and express them as functions of the elasticities and the elasticities of K_E, K and N with respect to taxes discussed in Section 4.1.

In the optimal tax problem, the first order condition for each $\tau_j \in \{\tau_K, \tau_W\}$ that ensures $\mathrm{d}\mathcal{W} = 0$ implies that

$$0 = B_{\tau_j} + \sum_{m \in \{K;W;N\}} \tau_m \frac{\partial B_{\tau_m}}{\partial \tau_j} - B_{\tau_j}^N N - \frac{(1-N)\, w\, (1-\tau_N)}{e_{\tilde{w}}^N} \frac{\partial N}{\partial \tau_j},$$

where $e_{\tilde{w}}^N = \frac{H'(z^*)}{N}$ and $B_{\tau_j}^N$ is the lifetime additional tax payments a worker would have to make after a unit increase in τ_j. The first term represents the additional tax revenue gained by increasing τ_j by one unit, while the term in the summation captures the (typically negative) revenue gains induced by the behavioral responses to the tax change. The third term represents the lifetime additional tax payments workers would have to make as increasing taxes requires adjusting τ_N to balance the government's budget. Lastly, the fourth term represents the resultant change in the relative welfare of entrepreneurs and workers, which, by revealed preference, can be inferred from the change in the number of workers after the tax. When taxes are set optimally, these effects balance out.

For each of the two first order conditions above, we can use the government budget constraint to eliminate τ_N and the definitions of the elasticities of tax bases with respect to taxes above to replace the $\frac{\partial B_{\tau_m}}{\partial \tau_j}$ terms. Rearranging, we obtain the vector of optimal taxes, summarized in the following proposition.

Proposition 1. *The optimal steady state tax vector* $\mathcal{T} = [\tau_K; \tau_W]^T$ *is given by*

$$\mathcal{T} = \left(A - g_1 + B^{-1}\left(-\mathcal{E} + \mathbf{e}^N \mathbf{1}^T\right) B\right)^{-1} \left(\mathbf{1} - \mathbf{g}_2 + B^{-1}\overline{G}\mathbf{e}^N\right), \tag{8}$$

where

$$\begin{aligned} g_1 &= AB^{-1}B^N - \left(e_{\tilde{w}}^N\right)^{-1}(1-N)\, B^{-1}\mathbf{e}^N \mathbf{1}^T B, \\ \mathbf{g}_2 &= \left(B^{-1}B^N\right)\mathbf{1} + \left(B_{\tau_N} - \overline{G}\right)\left(e_{\tilde{w}}^N\right)^{-1}(1-N)\, B^{-1}\mathbf{e}^N, \end{aligned}$$

and where $\mathbf{1}$ *denotes the column vector* $(1,1)^T$, $\mathbf{e}^N$ *denotes the column vector* $(e_{\tau_K}^N, e_{\tau_W}^N)^T$ *and*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} B_{\tau_K} & 0 \\ 0 & B_{\tau_W} \end{pmatrix}, \quad B^N = \begin{pmatrix} NB_{\tau_K}^N & 0 \\ 0 & NB_{\tau_W}^N \end{pmatrix}, \quad \mathcal{E} = \begin{pmatrix} e_{\tau_K}^{B_{\tau_K}} & e_{\tau_K}^{B_{\tau_W}} \\ e_{\tau_W}^{B_{\tau_K}} & e_{\tau_W}^{B_{\tau_W}} \end{pmatrix}.$$

Intuition and Implications. The formula abov bears noticeable similarities to scalar optimal linear tax formulae in the literature. These formulae often take the form $\tau = \frac{1-g}{1-g+|e|}$, where g is a function of marginal social welfare weights and e the elasticity of the tax base with respect to taxes (see, e.g. Saez and Stantcheva, 2018). In our case, the analogous term to $|e|$ is $-\mathcal{E} + \mathbf{e}^N \mathbf{1}^T$, which is a function of the elasticities of tax bases with respect to taxes, and the analogous terms to g are the matrix g_1 and vector $\mathbf{g}_2$. The formula reveals that optimal capital income and wealth taxes depend on three considerations: the extent to which they reduce the tax bases (the $-\mathcal{E}$ term); the extent to which they affect occupational choice and, therefore, labor income tax revenue (the $\mathbf{e}^N$ term), and the extent to which they fall on agents with a high marginal utility of consumption (the g_1 and $\mathbf{g}_2$ terms). Some of these effects are weighted by the size of tax bases and government spending. We next discuss these three considerations.

First, consistent with standard optimal tax principles, larger (negative) values of the elasticities in $\mathcal{E}$ make the optimal tax rates on capital and wealth smaller. That is, the optimal tax rates are lower if these taxes substantially reduce their tax bases, as Laffer curves peak at low tax rates. A consequence of the fact that the relevant elasticities in the optimal tax formula are partial equilibrium elasticities is that 'trickle down' effects of agents' choices on other agents via prices are irrelevant for optimal taxes. As such, the endogenous response of agents' behavior to tax changes *only* matters for optimal taxes insofar it affects the tax bases, and, therefore, tax revenue.

Second, $\mathbf{e}^N$ appears in the optimal tax formula because taxes affect the fraction of agents who are workers, which affects the labor income tax revenue. Across the various calibrations we study below, $\mathbf{e}^N$ is positive, indicating that higher capital income and wealth taxes increase the mass of workers and discourage entry into entrepreneurship. Provided $\overline{G}$ is sufficiently large, an increase in the magnitude of $\mathbf{e}^N$ will then, all else equal, increase optimal capital income and wealth taxes. This is because, for sufficiently high $\overline{G}$, government budget balance requires positive labor income taxes. In that case, a larger $\mathbf{e}^N$ implies that capital income and wealth taxes are effective at raising labor income tax revenue by discouraging entry into entrepreneurship.

Third, the matrix g_1 and the vector $\mathbf{g}_2$ are analogous terms to g in scalar linear optimal tax formulae and represent the direct effect of changes in τ_K and τ_W on social welfare. These terms depend on the extent to which taxes are paid by workers and affect entry into entrepreneurship. Intuitively, they can be interpreted as weights that the planner places on the payers of capital and wealth taxes, relative to the weight placed on the payers of labor taxes. The utilitarian planner places different weights on the payers of these taxes for insurance reasons: the payers of capital income taxes are frequently much richer, and so have a lower marginal utility of consumption on average, than the payers of labor income taxes.

The definitions of g_1 and $\mathbf{g}_2$ imply that if τ_K and τ_W were entirely paid by workers (and so tax changes did not directly affect entry into entrepreneurship) then $g_1 = A$, $\mathbf{g}_2 = \mathbf{1}$ and $\mathbf{e}^N = \mathbf{0}$ and, assuming the elasticities in $\mathcal{E}$ are positive, optimal taxes on capital income and wealth would be zero. Equally, the optimal tax formula reveals that *all* optimal taxes would similarly be zero if the government put equal weight on the consumption of the taxpayers of all three taxes (so that $\mathbf{g}_2 = \mathbf{1}$) and government spending $\overline{G}$ was equal to zero. This clarifies what the motives to tax capital income and wealth are: taxes on capital income and wealth are valuable for insurance reasons (i.e. if $\mathbf{g}_2 \neq \mathbf{1}$) and also valuable to influence entry into entrepreneurship (indicated by $\mathbf{e}^N$) which matters because it affects labor income tax revenue. As such, if there was no insurance motive to tax capital income and wealth and no government spending (and so no need for labor taxes) then all optimal tax rates would equal zero.

Alternatively, if the planner's objective was not to set capital income and wealth taxes to maximize welfare, but instead to minimize the value of τ_N, then repeating the steps above to derive the optimal tax formula implies that the planner's preferred tax vector $\mathcal{T}$ satisfies the same formula as in Proposition 1, except with $g_1 = 0$ and $\mathbf{g}_2 = 0$. This is roughly the tax vector that maximizes the combined steady state revenue from capital income and wealth taxes.[17] Intuitively, $g_1 = 0$ and $\mathbf{g}_2 = 0$ in this case because the planner is not concerned about the welfare of payers of capital income and wealth taxes. We consider this case in our calibration below.

Lastly, we note that the optimal tax formula obtained in Proposition 1 is not unique to our specific model but also holds in a relatively more general setting. In particular, our derivation of the optimal tax formula in Proposition 1 did not make use of many specific features of the model, including the specification of the financial friction, the functional form determining entrepreneurial risk, or the logarithmic utility. Therefore, irrespective of the assumptions on the utility function, entrepreneurial risk and agency frictions, we can express optimal taxes as a function of the size of the tax base for the various taxes, the degree to which each tax is borne by workers and entrepreneurs, and the partial equilibrium elasticities of the tax base with respect to each tax, according Proposition 1.[18] The specific assumptions we made on the utility function, entrepreneurial risk and agency frictions are necessary for characterizing the elasticities in closed form and for inferring the corresponding values for optimal taxes, but we are able to use the same formula when computing optimal taxes in a model with exogenous financial frictions.

[17] To be precise, since total government tax revenue must equal exogenous government spending $\bar{G}$ by budget balance, maximizing the revenue from wealth and capital income taxes is equivalent to minimizing the revenue from the labor income tax, $N \times \tau_N$. This is very close to minimizing τ_N. The two are not identical however, since N is endogenous in our model.

[18] See Appendix B.7 for a formal proof of this result.

5 Quantifying the Model

In this section we describe our calibration strategy and compute the optimal taxes implied by our numerical calibration of the economy outlined in Section 2.

5.1 Calibration Strategy

We set a number of parameters outside of the model and calibrate the rest to ensure our economy reproduces salient features of the US economy. We summarize the parameter values in Table 1.

Assigned Parameters. A period in the model is one year. Panel A of Table 1 summarizes the assigned parameters.

Demographics We set the mortality rate $\gamma = 2.5\%$, corresponding to a working life of 40 years, and set the depreciation rate $\delta = 0.07$, roughly the average depreciation rate in the US fixed asset tables.

Technology We set the distribution of the non-pecuniary taste for entrepreneurship z, to be an exponential distribution of the following form

$$H_z(z) = \begin{cases} 1 - h_0 e^{-\epsilon_E z} & \text{if } z > \frac{\log(h_0)}{\epsilon_E} \\ 0 & \text{otherwise.} \end{cases}$$

Given this functional form, ϵ_E can be interpreted as the elasticity of entry into entrepreneurship with respect to wages, and h_0 is the share of agents who actively enjoy entrepreneurship, that is have $z_i > 0$. As there are no direct estimates of the elasticity of entry into entrepreneurship with respect to wages, we set it equal to 1.5 as a baseline and discuss how optimal taxes vary with this parameter. We calibrate h_0 jointly with other parameters, as we discuss below. We set the distribution of θ, H_θ to be uniform on $[0, 1]$ and the autocorrelation of the productivity shock $1 - \lambda_\theta$ to 0.885, as in Cooper and Haltiwanger (2006).

Financial Frictions Since the entrepreneur's optimal contract is equivalent to an equity and debt contract, we set $\phi = 0.67$ to match the equity share of business owners in the US data. We use the Survey of Consumer Finances (National) Survey of Small Business Finances to document that entrepreneurs own, on average, 84% of their firm's equity.[19] Since $1 - \underline{\epsilon}$ represents the amount of within-period risk-free debt that entrepreneurs issue against their risky projects, as a share of project value, we choose $\underline{\epsilon}$ to match a debt-to-asset ratio for entrepreneurs of 0.35 (Crouzet and Mehrotra, 2020, Boar and Midrigan, 2019).

[19] See Appendix C for a detailed discussion of our treatment of the data.

Initial Tax System We abstract away from many details of the US tax system and assume linear tax rates on capital income and labor income (Floden and Lindé, 2001, Domeij and Heathcote, 2004, Heathcote, 2005, Dyrda and Pedroni, 2022, Guvenen et al., 2023) and a linear tax on wealth. We set the wealth tax in the initial steady state to zero, in line with current US tax policy, and calibrate distinct initial tax rates for capital and labor income. This captures the fact that, despite the US tax system applying a comprehensive income tax, in practice, the average effective rate of tax on capital income is different on from the average effective rate on labor income, because capital income is concentrated among high earners who pay higher marginal rates, and because some sources of capital income are subject to separate taxes (such as corporate tax).

The capital income tax in the model represents both the tax rate on the profits of entrepreneurs, and the tax rate on the entirety of the return to capital, since all capital is invested by entrepreneurs and so the total income derived from capital is equal to the profits of entrepreneurs. Both the average effective tax rate on capital income in the US and the tax rate on profits of non-corporate businesses (i.e. pass-through entities) appear close to 20% in recent times (McDaniel, 2007, CBO, 2014, Quantria-Strategies, 2009),[20] so we calibrate the initial capital income tax rate to 20%. We choose $\bar{G}$, so that the share of government spending is 20% of GDP and set τ_N so that the government's budget balances.[21]

Calibrated Parameters. Panel B of Table 1 summarizes the calibrated parameters. We assume that final output is produced with a Cobb-Douglas technology

$$Y = Y_E^{\alpha_E} Y_F^{\alpha_F} N^{1-\alpha_E-\alpha_F}.$$

We calibrate the technology parameters α_E and α_F together with the remaining parameters $\rho, \underline{k}_E, \varphi$, and h_0, which represent the discount rate, the level of capital that can be put into the risky technology without risk, the riskiness of the risky technology, and the fraction of households who enjoy entrepreneurship. We set these six parameters to match the following six steady state moments: a labor share of national income, the risk-free return and risky rate of return to capital net of depreciation, the capital-output ratio, the share of wealth held

[20] McDaniel (2007) calculates the average effective capital income tax rate, based on the tax revenue, to be 23% in 2003. She defines capital income taxes as encompassing taxes on corporate income, taxes paid by households on dividends, capital gains and on the capital share of income from private businesses, and property taxes paid by firms and other organizations. Based on a typical capital asset, CBO (2014) finds an average marginal effective rate of federal tax on capital income of 18% under 2014 law. Quantria-Strategies (2009) finds an average effective federal tax rate of for small businesses on their income, ranging from 13.3% for small non-farm sole proprieterships to 26.9% for small S corporations.

[21] We have experimented with varying the initial rate of capital income tax to 15% and 25%, recalibrating the other parameters. In both cases this shifts optimal capital and labor income tax rates by roughly 2 percentage points relative to the baseline and makes little difference to optimal wealth taxes.

by entrepreneurs, the share of households who are entrepreneurs. Lastly, as noted above, we set the labor income tax rate τ_N to ensure that the government's budget is balanced. We assume a labor share of 2/3 and a capital output ratio of 3, in line with the US national accounts. We use a risky return to capital (net of depreciation) of 8% and a risk-free return of 1%.[22] These are, respectively, the approximate average returns to equity and to relatively risk-free securities in the US over the twentieth century (Mehra and Prescott, 2003). We take the share of wealth held by entrepreneurs to be 53%, and the share of households who are entrepreneurs as 11.7%, as reported by Cagetti and De Nardi (2006) and Boar and Midrigan (2023), respectively, using data from the Survey of Consumer Finances.

Table 1: Parameter Values

Parameter	Value	Target moment
	Panel A. Assigned parameters	
γ	0.025	Working Life: 40 Years
δ	0.07	Depreciation
ϵ_E	1.5	Entry elasticity
λ_θ	0.115	Profitability autocorrelation
ϕ	0.67	Owner Equity Share (SSBF)
$\underline{\epsilon}$	0.350	Debt-to-asset ratio
τ_K	0.200	Corporate tax rate small businesses
τ_W	0	Current US level
$\bar{G}$	0.200	Government spending/GDP
	Panel B. Calibrated parameters	
α_E	0.188	Labor share 2/3
α_F	0.142	Risk-free rate
ρ	0.007	Capital-output ratio
$\underline{k}_E$	5.44	Entrepreneurs' share of wealth
φ	0.651	Return to Equity
h_0	0.176	Fraction of entrepreneurs
τ_N	0.263	Government budget balance

Untargetted Moments. One of the motives for taxation for the utilitarian planner is its desire to insure agents against idiosyncratic shocks. The strength of this motive is shaped by the amount of inequality in the economy resulting from these shocks. In our

[22] By choosing the parameters φ and $\underline{k}_E$ accordingly, the model is able to produce an arbitrarily large gap between risky and risk-free rates of return.

calibration, we only target the wealth share of entrepreneurs. However, our model is able to reproduce top wealth inequality more broadly. Table 2 compares the model's predictions regarding the wealth share held by the top 10%, 1% and 0.1% of the wealth distribution with the empirical shares reported by Piketty et al. (2018) and Smith et al. (2021). In both the data and the model, those at the top of the wealth distribution hold a large share of wealth, lending credibility to the model for studying the optimal capital and wealth taxation.

Table 2: Top Wealth Inequality

Wealth Share	Model	Data PSZ	Data SZZ
Top 10%	66.3%	73.4%	65.7%
Top 1%	41.0%	36.3%	31.5%
Top 0.1%	22.8%	18.4%	15.0%

Notes: The wealth shares in the column Data PSZ are from Piketty et al. (2018) and those in the column Data SZZ are from Smith et al. (2021).

Although not targeted, the model also produces an overall level and pattern of entrepreneurial risk that appears roughly in accordance with the data. We compare the level of risk faced by entrepreneurs in the initial steady state to the recent empirical findings of DeBacker et al. (2023). DeBacker et al. (2023) study a a panel of individual tax units over the period 1987-2018, and provide the standard deviation and mean business income per year. Taking the ratio of the two and averaging across years, we find that average coefficient of variation of business income is equal to 3.69. We simulate the model to produce a panel of 10,000 entrepreneurs, also for a 23 year period and find this ratio to be 3.54.[23] DeBacker et al. (2023) find business income to be positively skewed, with the average ratio of the median to mean at 0.26. In our panel, we also find it to be positively skewed, if a little less than in the data, with a ratio of median to mean at 0.58.

5.2 Implied Values of Elasticities

In Proposition 1 we showed that optimal taxes can be expressed in terms of tax bases and the elasticity of tax bases with respect to taxes. Before calculating optimal taxes, we first discuss the values our calibration implies for these determinants of optimal taxes. We focus our discussion on the term $\mathbf{g}_2$, and the elasticities of occupational choice and tax bases with

[23] For this exercise, we define the measured business income of an entrepreneur i as $\pi_{i,t}+(1-\delta)k_{i,t}-\hat{b}_{i,t}-a_{i,t}$. This implies that the entrepreneur's measured business income is (very slightly) affected by their choice of $k_{F,i,t}$, which is not pinned down in equilibrium since entrepreneurs are indifferent about their choice of $k_{F,i,t}$. We resolve this by assuming that all entrepreneurs choose the same value of $k_{F,i,t} - a_{i,t}$.

respect to taxes in the terms $\mathbf{e}^N$ and $\mathcal{E}$, and relegate the remaining terms to Appendix B.8. The calibrated values of the terms in the vectors $\mathbf{g}_2$, $\mathbf{e}^N$ and the matrix $\mathcal{E}$ are

$$\mathbf{g}_2 = \begin{pmatrix} 0.75 \\ 0.36 \end{pmatrix}, \quad \mathbf{e}^N = \begin{pmatrix} 0.29 \\ 0.19 \end{pmatrix}, \quad \mathcal{E} = \begin{pmatrix} -5.23 & -6.45 \\ -107.0 & -293.6 \end{pmatrix}.$$

Recall that the terms in the vector $\mathbf{g}_2$ roughly correspond to marginal social welfare weights on the tax payers of capital income and wealth taxes, relative to workers. Thus, the term 0.75 signifies that the planner puts substantial weight on the welfare of capital income tax payers, significantly more than the 0.36 weight it puts on the welfare of wealth tax payers. This is because a large fraction of capital income taxes are paid by relatively poor entrepreneurs, who earn high capital income relative to their wealth and have a high marginal utility of consumption. As we show below, this welfare weight on capital income tax payers reduces optimal capital income taxes and raises optimal wealth taxes.

The elasticity vector $\mathbf{e}^N$ summarizes the effect that capital income and wealth taxes have on the share of households who become workers. The vector contains elasticities slightly above zero, suggesting that increases in capital and wealth taxes mildly increase the number of agents who become workers, and thus have a relatively small effect on occupational choice.

The elasticity matrix $\mathcal{E}$, however, contains large negative elasticities. The diagonal terms indicate that a 1% increase in capital income tax rates reduces the capital income tax base by 5.2% and a 1% wealth tax reduces aggregate wealth by 293%. The latter number seems large, but since steady state rates of return to capital are small, even relatively low wealth taxes turn the rate of return to capital negative, severely weakening motivations to save. As such, to interpret the elasticity of wealth with respect to wealth taxes, it is instructive to multiply it by the risk-free rate of return. Let $e^{B_{\tau_W}}_{-\tilde{R}_F} \equiv \tilde{R}_F e^{B_{\tau_W}}_{\tau_W}$ denote the percentage decrease in wealth caused by a tax increase that reduces the risk-free rate by one percent of its initial value. In the calibrated steady state, we obtain $e^{B_{\tau_W}}_{-\tilde{R}_F} = -2.58$. While still relatively large, this is roughly comparable to Jakobsen et al. (2020), who use two Danish tax reforms to infer that a one percent decrease in the rate of return reduces aggregate wealth by 0.58-1.91% in the long run.[24] The reason for these relatively large effects, which provide a powerful motivation for the planner to set these taxes at relatively low levels, is that tax changes have large effects on K_E and K as small changes in the flow of household savings ultimately yield large long-run changes in the stock of aggregate wealth.

[24] We conjecture that the reason that our elasticity is still larger than the estimates of Jakobsen et al. (2020) is a consequence of a number of simplifying assumptions we made to keep the model parsimonious and tractable, including a simple log utility function, no habit persistence in preferences, no idiosyncratic risk for workers, and no retirement. Generalizing the model on these dimensions would presumably act to reduce the elasticity of wealth with respect to taxes.

5.3 Optimal Taxes

We report optimal taxes on capital income, wealth and labor income in Panel A of Table 3. In the first row of the table we calculate optimal tax rates by imputting the initial steady state values of the tax bases and elasticities into the formula in Proposition 1. In the second row, we acknowledge that this approach provides only an approximation for the optimal tax rates, as the size of tax bases and the elasticities of tax bases with respect to taxes can themselves change as taxes change. We therefore compute exact optimal tax rates by applying the optimal tax formula recursively. Specifically, we calculate (approximate) optimal taxes at the initial steady state, then recalculating the steady state at the new tax rates, and then recalculating (approximate) optimal taxes accordingly, repeating until convergence.

Whether approximate or exact, welfare maximization involves lower taxes on capital income than in the status quo, and a small positive tax on wealth. The optimal labor income tax is slightly higher than in the status quo. The largest difference between approximate and exact optimal taxes is in the case of the capital income tax: 8.3% vs. 3.7%. The reason for the difference is that cutting capital income taxes increases the elasticity of tax bases with respect to this tax compared to the initial steady state, encouraging the planner to move away from capital income taxes. We find that the consumption equivalent welfare gain from shifting from the status quo to the optimal welfare-maximizing taxes is 0.2%.[25]

That the approximate tax rates are comparable to the exact ones suggests that optimal taxes can be found by simply calculating the value of the elasticities in the status quo. This is useful for future research because, in principle, the elasticities in our optimal tax formula could be estimated empirically, making it possible to draw conclusions about optimal tax rates without having to commit to a particular model calibration.

To better understand the key factors determining these optimal tax rates, Panel B of the of the table shows the tax rates on capital and wealth that a planner would set absent the social welfare effects of taxes. This is identical to applying the formula in Proposition 1 while setting the terms in g_1 and $\mathbf{g}_2$ to zero. In doing so, we recover the capital income and wealth taxes that roughly maximize the revenue from these two taxes (i.e. specifically, minimize τ_N), as discussed above. We refer to these as revenue-maximizing taxes. The third line of the table shows the approximate revenue-maximizing taxes when the initial steady state elasticities are used, and the fourth line shows the exact revenue-maximizing taxes calculated recursively. Optimal revenue-maximizing taxes are much higher on capital income, and essentially zero on wealth. This reveals that the main reason for which it is

[25] As discussed above, the capital income tax rate can be interpreted as either a tax on the entirety of the return to capital, or a tax on entrepreneurial profits, since these are the same thing in the model. Therefore, our results are silent on whether it would be optimal to tax all capital income at a rate of 3.7% or whether a different tax rate should be applied to non-corporate businesses versus other forms of capital income.

Table 3: Optimal Taxes Under Welfare and Revenue Maximization

	τ_K^*	τ_W^*	τ_N^*
Panel A. Welfare Maximization			
Approximate optimal taxes	8.3%	0.1%	27.6%
Exact optimal taxes	3.7%	0.2%	28.0%
Panel B. Revenue Maximization			
Approximate optimal taxes	22.1%	0.0%	25.9%
Exact optimal taxes	20.4%	0.0%	26.2%

optimal to tax wealth and not solely capital income in the baseline model is that, as discussed above, the row of $\mathbf{g}_2$ corresponding to capital income taxes (0.75) is much higher than for wealth (0.36). Thus, capital income taxes are, in some respects, more undesirable than wealth taxes because of the significant negative welfare effect on poorer entrepreneurs.

5.3.1 Inspecting the Mechanism

To further understand the motivations behind capital income and wealth taxes in the model, we show how optimal taxes change as we vary the parameter ϕ, which governs the severity of financial frictions in Figure 1, $\underline{k}_E$ which governs the return to scale for entrepreneurs putting capital into the risky technology in Figure 2, and ϵ_E which determines the elasticity of entry intro entrepreneurship with respect to taxes in Figure 3. The top row of each figure shows the welfare and revenue maximizing tax rates and the bottom row shows the elasticities of capital income (with respect to τ_K), wealth (with respect to the post-tax rate of return) and N (with respect to τ_K) at the initial steady state and at the optimal tax rates.

A unifying theme that emerges from these figures is that optimal wealth taxes are decreasing in ϕ, $\underline{k}_E$ and ϵ_E, whereas optimal capital income taxes are increasing in these parameters, and can even be negative if the values of the parameters are low enough. The vertical line marks the value of each parameter in our benchmark calibration.

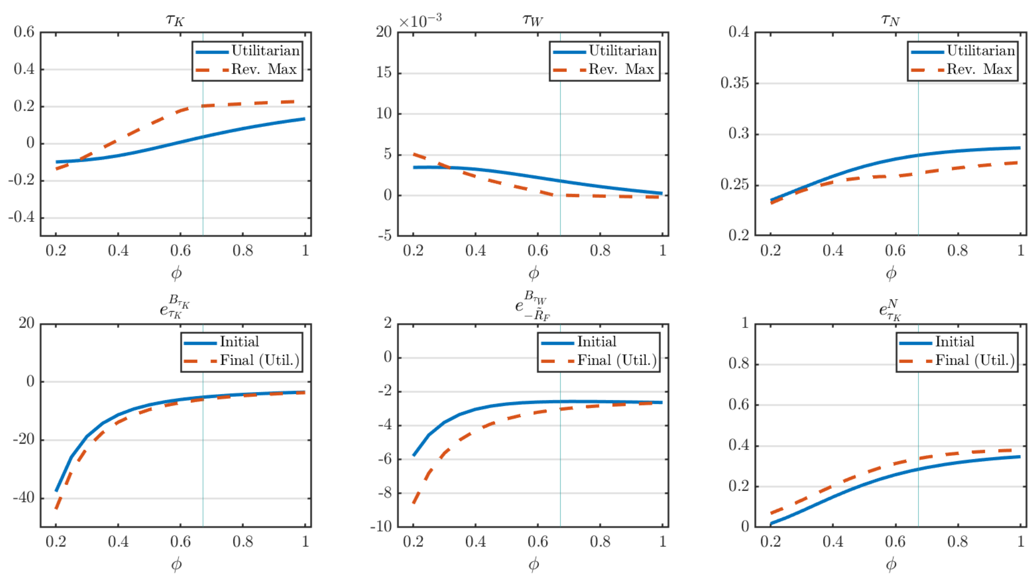
To understand the intuition behind the effects at play, recall from Section 4.1 that the elasticity of risky capital with respect to taxes is

$$e_{\tau_j}^{K_E} \propto \frac{(1-N)\underline{k}_E}{K_E},\ j \in \{K; W\}\,.$$

Here, the right hand side represents the part of K_E that arises from each entrepreneur, inelastically, allocating $\underline{k}_E$ to the risky technology and is, therefore, a determinant of the elasticity of capital income with respect to taxes. When this is close to 1, all entrepreneurs put roughly $\underline{k}_E$ into the risky technology, so capital income is inelastic in response to taxation, whereas when it is small, capital income is potentially responsive to taxation. In this case, changes in capital income taxes have a bigger effect on tax bases than similarly fiscally large changes in wealth taxes, because capital income taxes also affect the excess return $(r_E - r_F)(1 - \tau_K)$ that determines the allocation capital to the risky technology, whereas wealth taxes do not.[26] Therefore, the parameters ϵ_E, $\underline{k}_E$ and ϕ affect optimal taxes through their effect on $\frac{(1-N)\underline{k}_E}{K_E}$.

Consider first the effect of ϕ and $\underline{k}_E$, depicted in Figures 1 and 2. Higher ϕ and $\underline{k}_E$ both increase $\frac{(1-N)\underline{k}_E}{K_E}$, rendering capital income inelastic. Since in this case a large share of the capital income tax falls on the inelastic $\underline{k}_E$ capital in the risky technology, capital income taxes can raise more revenue relative to their effects on aggregate saving than wealth taxes.

Figure 1: Optimal Taxes As Financial Frictions Vary

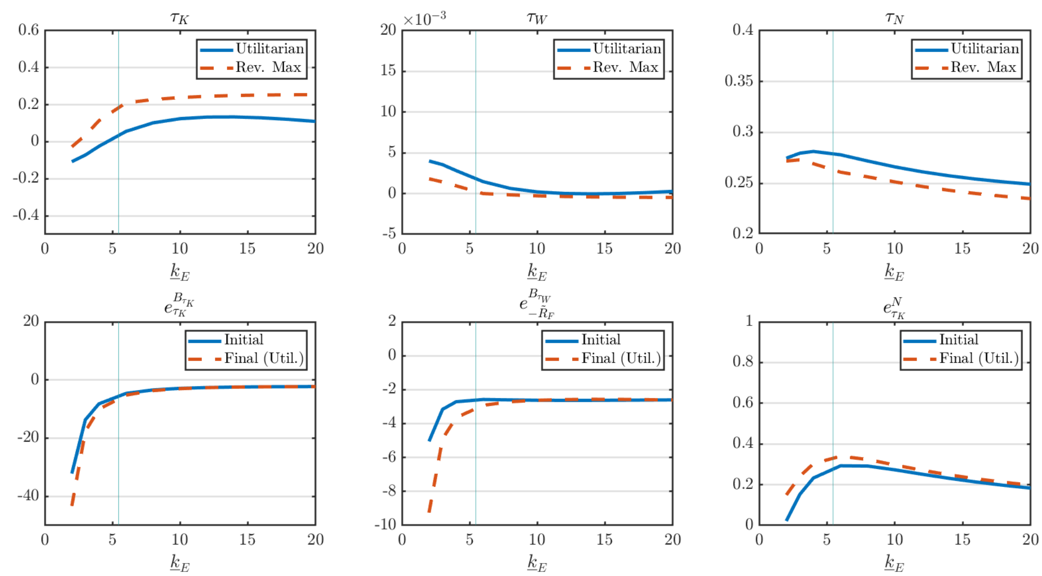


Notes: The vertical line denotes the benchmark value of ϕ. The blue line on the bottom row shows the elasticities at the initial steady state, and the red line shows the elasticities at the optimal tax rates.

Turning to the effects of ϵ_E, depicted in Figure 3, we find that a high value of ϵ_E also motivates a higher rate of capital income taxes and a lower rate of wealth taxes. This is because an additional effect of low capital income taxes is that they encourage entry into entrepreneurship, which reduces overall tax revenue if workers pay more tax than entrepreneurs. Wealth taxes do not affect entry into entrepreneurship in the same way, since workers and entrepreneurs are similarly affected by these taxes, given household wealth. When $\epsilon_E \to 0$

[26] This is very similar to the 'use it or lose it' argument for taxing wealth in Guvenen et al. (2023).

Figure 2: Optimal Taxes As Returns to Scale Vary

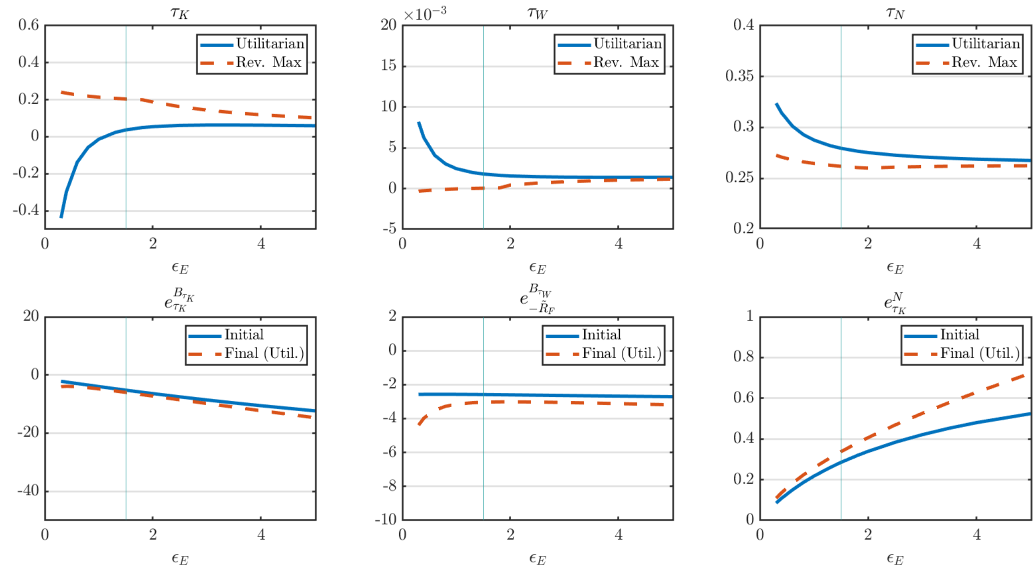


Notes: The vertical line denotes the benchmark value of $\underline{k}_E$. The blue line on the bottom row shows the elasticities at the initial steady state, and the red line shows the elasticities at the optimal tax rates.

and entry is inelastic, it is optimal to tax wealth and subsidize capital income, as in Guvenen et al. (2023). Although we modeled entry into entrepreneurship in a parsimonious way so as to preserve analytical tractability, our results suggest the importance of this margin for determining the relative merits of capital income and wealth taxes and the need for further empirical evidence to discipline this margin.

In sum, this discussion suggests that the choice of optimal tax rates on capital income and wealth involves a tradeoff between three competing motivations. First, higher capital income or wealth tax rates mechanically raises tax revenue, especially from richer entrepreneurs, which can be passed to workers as lower labor taxes, which is desirable to a utilitarian planner. Since capital income taxes partly fall on relatively poor entrepreneurs, an insurance motive favors taxing capital income less and wealth more. Second, higher tax rates on capital income and wealth discourage saving and investment in the high-return risky technology. The effect of capital income taxes on the allocation of capital motivates low capital income taxes and higher wealth taxes. This is mitigated by the fact that a part of capital income is relatively insensitive with respect to taxes, which motivates higher taxes on capital income and lower taxes on wealth. Third, higher capital income tax rates, in particular, discourage entry into entrepreneurship. If workers pay a higher total tax per capita than entrepreneurs, this increases tax revenue, which can then be passed on to workers as lower labor tax rates.

Figure 3: Optimal Taxes As Entry Elasticity Varies



Notes: The vertical line denotes the benchmark value of ϵ_E. The blue line on the bottom row shows the elasticities at the initial steady state, and the red line shows the elasticities at the optimal tax rates.

5.3.2 Optimal Taxes and the Nature of Financial Frictions

Throughout the paper, we have assumed that financial frictions arise endogenously as a consequence of the private information entrepreneurs have about their idiosyncratic shocks. As discussed, a feature of this environment is that the ability of entrepreneurs to obtain external finance for risky capital investment evolves endogenously in response to tax changes. This is distinct from what is most commonly assumed in the literature, including the recent literature on optimal taxation with financial frictions (e.g. Guvenen et al., 2023, Panousi, 2012, Boar and Midrigan, 2023), where entrepreneurs' ability to obtain external equity or external finance is constrained to be an exogenous fraction of their net worth.

We next show that the endogeneity of financial frictions to taxes matters for the design of optimal taxes. To that end, we compare our endogenous financial friction to two exogenous frictions typically used in the literature. We show that all three frictions yield different results, indicating that microfounding financial frictions is important for understanding the consequence of these frictions for optimal capital taxation and simply taking the nature of the frictions to be exogenous could produce misleading inferences about optimal taxation.

The first exogenous friction we consider is similar to the collateral constraints typically assumed in the literature (e.g. Cagetti and De Nardi, 2006, Buera et al., 2011, Midrigan and Xu, 2014, Guvenen et al., 2023, Boar and Midrigan, 2023). In particular, we assume that an entrepreneur's choice of risky capital investment, k_E is constrained to satisfy

$$k_E \leq \underline{k}_E + \lambda\theta P,$$

where λ is an exogenous parameter. We assume that entrepreneurs face no uninsurable ex-post idiosyncratic risk (i.e. the variance of ξ is set to zero), which ensures that this financial constraint binds with equality in equilibrium provided $r_E > r_F$. We calibrate λ and recalibrate the mean of the preference shock z_i to target the same risky rate of return and entry as in the baseline model, so that all variables take the same values in the benchmark.[27]

The second exogenous financial friction we consider is consistent Panousi and Reis (2014). In this case, we assume that entrepreneurs face idiosyncratic risk as in the baseline, but the fraction of external equity they can sell is exogenously equal to a constant λ. This creates an exogenous amount of undiversifiable idiosyncratic entrepreneurial risk, as in Panousi and Reis (2014). In our baseline, where financial frictions arise endogenously from information frictions, the equilibrium contract is one where the fraction of equity entrepreneurs can sell is endogenous and is decreasing in the capital income tax, as shown in Lemma 1. We calibrate λ in this alternative financial friction model to match the same level of external equity issuance as in the baseline model, so that all variables take the same values in the benchmark.

Panel A of Table 4 reports the value of optimal taxes in the economies where entrepreneurs are subject to these two alternative exogenous financial frictions. The first column of the table reproduces the values of optimal taxes from the baseline with endogenous financial frictions. The three models make different predictions regarding the optimal rate of capital income taxation. The optimal capital income tax rate is 3.7% in our baseline, and is much higher in the two models with exogenous financial frictions: 8.3% in the model with collateral constraints and as high as 25% in the exogenous equity share model. By contrast, the level of optimal wealth taxes is relatively similar across models.

To help interpret the difference in the magnitude of optimal taxes across models, Panel B reports the values of several of the key elasticities driving optimal taxation choices. As the panel shows, the elasticity of wealth with respect to wealth taxes, $e^{B_{\tau_W}}_{-\tilde{R}_F}$, is unaffected by the nature of the financial friction, so the three models prescribe similar levels for the wealth tax. This is because the financial friction does not affect agents' tradeoff between consuming and saving in the risk-free technology.

However, the nature of the financial friction makes a substantial difference to the elasticity of capital income with respect to capital income taxes, $e^{B_{\tau_K}}_{\tau_K}$, which is much larger in the baseline model than with either of the two alternative exogenous financial frictions. The reason for this is apparent from the expression for $e^{K_E}_{\tau_j}$ in Section 4.1: the elasticity

[27] An alternative form of the financial constraint we have considered is $k_E \leq \underline{k}_E + \lambda P$. We find that this has almost identical implications for optimal taxes.

Table 4: Model Results with Exogenous Financial Frictions

	Baseline	Collateral Constraint	Exogenous Equity Share
		Panel A. Optimal Taxes	
τ_K	3.7%	8.3%	24.6%
τ_W	0.18%	0.11%	-0.13%
τ_N	28.0%	27.6%	26.1%
		Panel B. Initial Elasticities	
$e_{\tau_K}^{B_{\tau_K}}$	-5.23	-4.13	-2.53
$e_{-\tilde{R}_F}^{B_{\tau_W}}$	-2.58	-2.58	-2.58

Notes: The optimal tax rates in Panel A are those that maximize long-run utilitarian welfare.

$e_{\tau_K}^{\hat{k}}$, reflecting the average willingness of entrepreneurs to invest capital into the risky technology, is approximately -1 in the baseline model, compared to either close to 0 or positive with exogenous financial frictions. This is because, in the baseline model, as discussed in Section 3.2, an increase in capital income taxes reduces the post-tax excess return from risky projects, but does not reduce the risk from these projects to the entrepreneur under the equilibrium financial contract, and so these taxes discourage risky investment. With the alternative financial frictions, this effect does not arise. In the case of the exogenous collateral constraint, the level of risky capital investment, conditional on present-value net worth and entrepreneurial ability, is unaffected by τ_K. In the case of the exogenous equity share model, a higher τ_K in fact encourages investment in risky projects because the tax on capital gains reduces the variance of entrepreneurial profits and so reduces the idiosyncratic risk from risky investment.[28] In the baseline model, the equilibrium share of equity the entrepreneur can sell externally decreases as τ_K rises, but this share does not not decrease in the exogenous equity constraint model.

These results suggest that the precise form of the financial friction, and its endogenous evolution in response to taxes, is important in assessing the elasticity of capital income with respect to capital income taxes, and therefore for the level of optimal capital taxation. This also points towards the potential of empirical estimates of this elasticity to serve as a disciplining device for models of financial frictions.

[28] In the exogenous equity share model, $\hat{k}_E(\theta) = \frac{1}{\phi(1-\underline{\epsilon})(1-\tau_K)} \times \min\left[\frac{(r_E - r_F)\theta}{\phi(1-\underline{\epsilon})(1-\tau_K)\varphi^2}; 1\right]$, which is identical to the expression for $\hat{k}_E(\theta)$ from the baseline, except for the additional $1-\tau_K$ term.

5.3.3 Sensitivity analysis

Entrepreneurial ability. We explore how our results depend on other features of the environment. As is well known in the public finance literature, optimal taxes typically depend on the underlying distribution of ability. We find this matters little for the results. To illustrate, we consider two alternative calibrations of the autocorrelation of entrepreneurial ability, which is governed by the parameter λ_θ. In the first calibration we reduce λ_θ from 0.115 to 0.05, corresponding to a rather persistent process for entrepreneurial ability. In the second one, we increase this parameter to 0.5, indicative of little persistence in entrepreneurial ability. Optimal tax rates are almost identical to our baseline model (within one percentage point for each tax rate) in each of these cases, consistent with our result that, provided λ_θ is high enough, the elasticities $e^{\hat{k}_E}_{\tau_K}$ and $e^{\hat{k}_E}_{\tau_W}$ are close to -1 and 0, regardless of the exact level of λ_θ. It turns out that even a λ_θ of 0.05 is high enough for the result to hold, since it generates a low correlation between wealth and ability, similer to the case of a λ_θ being iid.

Selection into entrepreneurship. Throughout, we have assumed that entry into entrepreneurship depends on the realization of an idiosyncractic taste shock z_i, so that there is no selection into entrepreneurship on the basis of entrepreneurial productivity or wealth. If such effects were allowed, for example by assuming that agents draw a signal about their entrepreneurial ability θ_i at birth, this would tend to reduce the elasticity of capital income with respect to capital income taxes. This is because higher capital income tax rates tend to discourage agents from becoming entrepreneurs, increasing the tendency to which only agents with a high θ_i become entrepreneurs. Therefore, average entrepreneurial ability will be increasing in the capital income tax rate. Since high ability entrepreneurs invest relatively more in the risky technology and earn relatively more capital income, this limits the extent to which the capital income tax reduces the tax base.

Labor supply. We have also assumed that workers supply labor inelastically and so taxes only affect labor supply at the extensive margin by affecting agents' occupational choice. In a previous version of the paper, we have assumed that workers have GHH preferences, which capture substitution but not income effects of labor income taxes on hours worked. In that case, we derived an optimal tax formula that closely resembles the optimal tax formula in Proposition 1. The only distinction is that the elasticities of the mass of workers with respect to taxes $\mathbf{e}^N$ in equation (8) are replaced by the elasticities of aggregate labor income with respect to taxes.[29] Therefore, our discussion on the intutition and implications of the optimal tax formula in Section 4.3 applies. Outside of GHH preferences, we can no

[29] The g_1 and $\mathbf{g}_2$ terms, which also involve the elasticities $\mathbf{e}^N$, are exactly as in Proposition 1.

longer easily derive a formula for optimal taxes that allows us to isolate the determinants of optimal taxes, but we believe that the insights of our analysis would be unchanged.

6 Conclusion

We examine optimal linear taxation in a setting with endogenous entry and financial frictions. Financial frictions imply that the distribution of wealth across entrepreneurs with different ability levels affects how efficiently capital is allocated in the economy – a force missing from models without financial frictions. That financial frictions are endogenous implies that taxes affect the allocative efficienty of capital. The planner chooses taxes on capital income, wealth and labor income to maximize the steady state welfare of a newborn agent. In the model, newborn agents decide whether to become workers or entrepreneurs. Workers supply labor inelastically, while entrepreneurs operate a production technology that uses capital and are subject to a financial constraint. As in the data, entrepreneurs are relatively richer on average, leading to an equity motive for capital income and wealth taxation.

Our model is analytically tractable and we characterize optimal steady state taxes as closed-form functions of the size of tax bases and the elasticity of tax bases with respect to taxes, in the tradition of the 'sufficient statistics' approach to optimal taxation. When we calibrate the model, we find that it is optimal to tax both capital income and wealth at relatively low but positive rates. We find that modelling financial frictions endogenously is consequential for optimal taxation. Our level of optimal capital income tax is lower than in otherwise identical models with exogenous financial frictions.

References

Abo-Zaid, S. (2014). Optimal capital-income taxation in a model with credit frictions. *The B.E. Journal of Macroeconomics*, 14(1):147 – 172.

Basseto, M. and Cui, W. (2020). A ramsey theory of financial distortions. Working paper.

Benhabib, J. and Szőke, B. (2021). Optimal positive capital taxes at interior steady states. *American Economic Journal: Macroeconomics*, 13(1):114–50.

Biljanovska, N. (2019). Optimal Policy In Collateral Constrained Economies. *Macroeconomic Dynamics*, 23(2):798–836.

Biljanovska, N. and Vardoulakis, A. P. (2019). Capital taxation with heterogeneous discounting and collateralized borrowing. *Journal of Macroeconomics*, 60:97 – 109.

Boar, C. and Midrigan, V. (2019). Markups and inequality. Working Paper 25952, National Bureau of Economic Research.

Boar, C. and Midrigan, V. (2020). Efficient redistribution. Working Paper 27622, National Bureau of Economic Research.

Boar, C. and Midrigan, V. (2023). Should we tax capital income or wealth? *American Economic Review: Insights*, 5(2):259–74.

Brüggemann, B. (2021). Higher taxes at the top: The role of entrepreneurs. *American Economic Journal: Macroeconomics*, 13(3):1–36.

Buera, F. J., Kaboski, J. P., and Shin, Y. (2011). Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002.

Cagetti, M. and De Nardi, M. (2006). Entrepreneurship, frictions, and wealth. *Journal of political Economy*, 114(5):835–870.

Cagetti, M. and De Nardi, M. (2009). Estate taxation, entrepreneurship, and wealth. *American Economic Review*, 99(1):85–111.

CBO (2014). Taxing capital income: Effective marginal tax rates under 2014 law and selected policy options. Technical report, Congress of the United States Congressional Budget Office.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54(3):607–622.

Chari, V. V. and Kehoe, P. J. (1999). Optimal fiscal and monetary policy. Working Paper 6891, National Bureau of Economic Research.

Chen, Y., Chen, Y., and Yang, C. (2019). Implementing the modified golden rule? optimal ramsey taxation with incomplete markets revisited. Working paper.

Cooper, R. W. and Haltiwanger, J. C. (2006). On the Nature of Capital Adjustment Costs. *The Review of Economic Studies*, 73(3):611–633.

Crouzet, N. and Mehrotra, N. R. (2020). Small and large firms over the business cycle. *American Economic Review*, 110(11):3549–3601.

Cullen, J. B. and Gordon, R. H. (2007). Taxes and entrepreneurial risk-taking: Theory and evidence for the us. *Journal of Public Economics*, 91(7-8):1479–1505.

Dávila, E. and Hébert, B. (2023). Optimal corporate taxation under financial frictions. *The Review of Economic Studies*, 90(4):1893–1933.

DeBacker, J., Panousi, V., and Ramnath, S. (2023). A risky venture: Income dynamics among pass-through business owners. *American Economic Journal: Macroeconomics*, 15(1):444–74.

Devereux, M. (2009). Taxing risky investment. Technical report, Oxford University Centre for Business Taxation.

Diamond, P. A. and Mirrlees, J. A. (1971). Optimal taxation and public production i: Production efficiency. *The American Economic Review*, 61(1):8–27.

Domeij, D. and Heathcote, J. (2004). On the distributional effects of reducing capital taxes. *International Economic Review*, 45(2):523–554.

Dyrda, S. and Pedroni, M. (2022). Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk. *The Review of Economic Studies*, 90(2):744–780.

Evans, D. (2015). Optimal taxation with persistent idiosyncratic investment risk. Working paper.

Floden, M. and Lindé, J. (2001). Idiosyncratic risk in the united states and sweden: Is there a role for government insurance? *Review of Economic Dynamics*, 4(2):406–437.

Guvenen, F., Kambourov, G., Kuruscu, B., Ocampo, S., and Chen, D. (2023). Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation*. *The Quarterly Journal of Economics*, 138(2):835–894.

Hayek, F. A. (1960). *The Constitution of Liberty.(Second Impression.)*. University of Chicago Press.

Heathcote, J. (2005). Fiscal Policy with Heterogeneous Agents and Incomplete Markets. *The Review of Economic Studies*, 72(1):161–188.

Hurst, E. and Pugsley, B. W. (2011). What do small businesses do? *Brookings Papers on Economic Activity*, pages 73–143.

Itskhoki, O. and Moll, B. (2019). Optimal development policies with financial frictions. *Econometrica*, 87(1):139–173.

Jakobsen, K., Jakobsen, K., Kleven, H., and Zucman, G. (2020). Wealth taxation and wealth accumulation: Theory and evidence from denmark. *The Quarterly Journal of Economics*, 135(1):329–388.

Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1):59 – 83.

Kitao, S. (2008). Entrepreneurship, taxation and capital investment. *Review of Economic Dynamics*, 11(1):44–69.

McDaniel, C. (2007). Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003. *Manuscript, Arizona State University*, 19602004.

Mehra, R. and Prescott, E. C. (2003). The equity premium in retrospect. *Handbook of the Economics of Finance*, 1:889–938.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–58.

Panousi, V. (2012). Capital taxation with entrepreneurial risk.

Panousi, V. and Reis, C. (2014). Optimal capital taxation with idiosyncratic investment risk. Working paper.

Panousi, V. and Reis, C. (2019). A unified framework for optimal taxation with undiversifiable risk. Working paper.

Phelan, T. (2019). On the Optimality of Differential Asset Taxation. Working papers, Federal Reserve Bank of Cleveland.

Piketty, T. and Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, 81(5):1851–1886.

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional national accounts: methods and estimates for the united states. *The Quarterly Journal of Economics*, 133(2):553–609.

Quantria-Strategies (2009). Effective Federal income tax rates faced by small businesses in the United States. Technical Report 343, Small Business Administration, Office of Advocacy, Quantria Strategies, LLC, Cheverly, MD 20785.

Rotberg, S. and Steinberg, J. B. (2019). Tax the wealthy? a dynamic general equilibrium analysis of progressive wealth taxation. Working paper.

Saez, E. and Stantcheva, S. (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162:120 – 142.

Shourideh, A. (2014). Optimal Taxation of Wealthy Individuals. Working paper, Wharton School.

Smith, M., Zidar, O. M., and Zwick, E. (2021). Top wealth in america: New estimates and implications for taxing the rich. Technical report, National Bureau of Economic Research.

Straub, L. and Werning, I. (2020). Positive long-run capital taxation: Chamley-judd revisited. *American Economic Review*, 110(1):86–119.

Appendices

A Discrete Time Model

A.1 Worker's Optimization Problem

We derive the solution to the worker's problem. Let

$$P_t^N \equiv a_t^N + \underbrace{\sum_{j=0}^{\infty} \left[\frac{w_{t+j}(1-\tau_{N,t+j})(1-\gamma)^j}{\Pi_{k=0}^{j} R_{F,t+k}} \right]}_{F_t^N}$$

denote the discounted value of the worker's lifetime income. Abusing notation, the worker problem can be written recursively as

$$V^N(P^N, X) = \max_{c \geq 0, P' \geq 0} \left(\log\left(c^N\right) + (1-\rho)\,(1-\gamma)\, V^N\left(P^{N'}, X'\right)\right)$$

$$\text{s.t.} \quad c^N + (1-\gamma)P^{N'} = R_F P^N$$

We solve this problem by *guess and verify*. We guess

$$V^N(P^N, X) = Q + \frac{1}{1-(1-\rho)(1-\gamma)} \log\left(P_t^N\right),$$

take first order conditions and use the guess to obtain the worker's policy functions

$$P^{N'} = (1-\rho)\,(1-\gamma)\,\frac{R_F}{1-\gamma} P^N \quad \text{and} \quad c^N = \left[1-(1-\rho)\,(1-\gamma)\right] R_F P^N.$$

To verify the guess, we plug the policy functions back into the value function and rearrange to obtain $V^N(P^N, X) = \underbrace{V^N(1, X)}_{=Q} + \frac{1}{1-(1-\rho)(1-\gamma)} \log P^N$. □

A.2 Entrepreneur's Optimization Problem

Let $P_{i,t}$ be the present value of the resources obtained by an entrepreneur who takes no risk: she puts exactly $\underline{k}_E$ units of capital into the risky technology, no capital in the risk-free technology, and lends the remaining $a_{i,t} - \underline{k}_E$ to banks at rate $R_{F,t}$. Therefore, $P_{i,t}$ is

$$P_{i,t} = a_{i,t} + \sum_{j=0}^{\infty} \frac{[1 + (r_{E,t} - \delta)(1-\tau_{K,t+j}) - \tau_{W,t} - R_{F,t}]\underline{k}_E (1-\gamma)^j}{\Pi_{k=0}^{j} R_{F,t+k}} \equiv a_{i_t} + F_t.$$

The morning and evening budget constraints can then be rewritten as $b_{i,t} = k_{E,i,t} + k_{F,i,t} - P_{i,t} + F_t$ and $c_{i,t} + (1-\gamma) P_{i,t+1} = \omega_{i,t}$, where end-of-period lifetime resources $\omega_{i,t}$ satisfy

$$\begin{aligned} \omega_{i,t} &= [\phi - (1-\tau_{K,t})r_{E,i,t} - (1-\delta)(1-\tau_{K,t})]k_{H,i,t} - (\hat{b}_{i,t} - R_{F,t}b_{i,t}) + R_{F,t}P_{i,t} \\ &+ [(1-\tau_{K,t})\, r_{E,t} + (1-\delta)(1-\tau_{K,t})](\tilde{k}_{E,i,t} - \underline{k}_E) + (\tau_{K,t} - \tau_{W,t} - R_{F,t})\,(k_{E,i,t} - \underline{k}_E) \\ &+ [-R_{F,t} + 1 + (1-\tau_{K,t})\,(r_{F,t} - \delta) - \tau_{W,t}]k_{F,i,t} \end{aligned}$$

Letting $\tilde{V}(\omega, \theta, X)$ denote the value in the evening of an entrepreneur with lifetime resources ω, we can write the entrepreneur's between period problem recursively as

$$\tilde{V}(\omega, \theta, X) = \sup_{c,P'} \log(c) + (1-\rho)\,(1-\gamma)\mathbb{E}V(P', \theta', X'), \tag{A.1}$$

$$\text{s.t.} \quad c + (1-\gamma)\, P' = \omega. \tag{A.2}$$

The entrepreneur's within-period problem is

$$V(P, \theta, X) = \sup \int_\xi \tilde{V}(\omega, \theta, X) \mathrm{d}H_\xi(\xi),$$

$$\text{s.t.} \quad b = k_E + k_F - P + F \tag{A.3}$$

$$0 = \int_\xi (\hat{b} - R_F b)\mathrm{d}H_\xi(\xi) \tag{A.4}$$

$$\begin{aligned} \omega &= [\phi - (1-\tau_K)r_E - (1-\delta)(1-\tau_K)]k_H - (\hat{b} - R_F b) + R_F P \\ &+[(1-\tau_K)\, r_E + (1-\delta)(1-\tau_K)](\tilde{k}_E - \underline{k}_E) + (\tau_K - \tau_W - R_F)\,(k_E - \underline{k}_E) \\ &+[-R_F + 1 + (1-\tau_K)\,(r_F - \delta) - \tau_W]k_F \end{aligned} \tag{A.5}$$

$$k_H \leq k_E \tag{A.6}$$

$$\frac{\partial \omega}{\partial \xi} \geq \phi \frac{\partial \tilde{k}_E}{\partial \xi}. \tag{A.7}$$

Our assumptions on the function q imply that $\tilde{k}_E$ depends linearly on k_E and so ω depends linearly on k_E and k_F. Together with log utility, this implies that the value functions $V(P, \theta, X)$ and $\tilde{V}(\omega, \theta, X)$ are equal to

$$\begin{aligned} V(P, \theta, X) &= \overline{V}(\theta, X) + \frac{\log(P)}{1 - (1-\rho)(1-\gamma)}, \\ \tilde{V}(\omega, \theta, X) &= \tilde{V}(1, \theta, X) + \frac{\log(\omega)}{1 - (1-\rho)(1-\gamma)}, \end{aligned}$$

where $\overline{V}(\theta, X) = V(1, \theta, X)$. The solution to the between period problem is then easily shown to be $c = (1 - (1-\rho)\,(1-\gamma))\,\omega$ and $P' = (1-\rho)\,\omega$.

As noted in the text, the incentive compatibility constraint (A.7) can be written as

$$\omega(P,\theta,\xi,X) \equiv \underline{\omega}(P,\theta,X) + \phi(q(\theta,\xi,k_E(P,\theta,X)) - \underline{k}_E).$$

Combining it with the definition of ω, integrating with respect to ξ and using equations (5)-(6) reveals that $\underline{\omega}$ must satisfy

$$\underline{\omega} + \phi(k_E - \underline{k}_E) = R_F P + (1-\tau_K)(r_E - r_F)(k_E - \underline{k}_E).$$

Then, we can rewrite the entrepreneur's within-period problem more compactly in terms of choosing functions $k_E(P,\theta,X) \geq \underline{k}_E$ and $\underline{\omega}(P,\theta,X)$ to solve

$$\sup \int_\xi \log\left(\underline{\omega} + \phi(\tilde{k}_E - \underline{k}_E)\right) \mathrm{d}H_\xi(\xi),$$

$$\begin{aligned} \text{s.t.} \quad \underline{\omega} &= \left(-\phi + (r_E - r_F)(1-\tau_K)\right)(k_E - \underline{k}_E) + R_F P \\ 0 &\leq \underline{\omega} + \phi(k_E - \underline{k}_E), \end{aligned}$$

and where $\tilde{k}_E = q(\theta,\xi,k_E)$. This is a standard portfolio choice problem, where there is a trade-off between risk and return. Choosing a higher k_E increases the variance of $\tilde{k}_E$ and therefore of ω, since $\omega = \underline{\omega} + \phi\tilde{k}_E$, but a higher k_E will increase the expected value of ω.

A.3 Proof of Lemma 1

We first evaluate the end of period entrepreneur lifetime resources ω in a contract where the entrepreneur issues debt and equity as indicated in the Lemma and show that it the same value of ω in every state of the world as the equilibrium contract.

Suppose that the entrepreneur issues risk-free debt $\tilde{b}$ by leveraging her risky and risk-free projects and sells fraction s of the leveraged value of her projects as equity, where

$$s = 1 - \frac{\phi}{(1-\tau_K)(r_E + 1 - \delta)} \tag{A.8}$$

We allow for the possibility that $\tilde{b} < 0$ in which case the entrepreneur is lending to the bank.

Let V_P denote the value of the entrepreneur's projects at the end of the period, gross of debts. If the entrepreneur did not interact with financial markets at all (i.e. set b and $\hat{b} = 0$ in every state), then V_P would be the same as the entrepreneur's end of period resources ω. Then, equation (A.5) implies that V_P is given by $V_P = \omega + \hat{b}$, where $\hat{b}$ is the total payout by the entrepreneur at the end of the period associated with the debt and equity contract.

Equivalently, equation (A.5) and (5) imply that

$$V_P = R_F P + R_F b + (1-\tau_K)(r_E + 1 - \delta)(\tilde{k}_E - \underline{k}_E) + (\tau_K - \tau_W - R_F)(k_E - \underline{k}_E). \quad \text{(A.9)}$$

The total external funds the entrepreneur can obtain at the start of the period, b, are

$$b = \tilde{b} + \frac{s(E_\xi[V_P] - R_F\tilde{b})}{R_F} \quad \text{(A.10)}$$

where $V_P - R_F\tilde{b}$ is the levered end-of-period value of the entrepreneur's projects, and $\frac{s(V_P - R_F\tilde{b})}{R_F}$ is the amount that the financial intermediary would be willing to pay for fraction s of the equity in these projects, given that it must earn the risk-free rate in expectation.

The entrepreneur who sells fraction s of the equity in her projects and borrows amount $\tilde{b}$ against these projects will have end of period resources

$$\omega = (1-s)(V_P - R_F\tilde{b}). \quad \text{(A.11)}$$

Substituting in equations (A.8), (A.9) and (A.10) and using that $E_\xi\tilde{k}_E = k_E$, we obtain

$$\omega = R_F P + \phi(\tilde{k}_E - \underline{k}_E) + [(1-\tau_K)(r_E - \delta) + 1 - \tau_W - R_F - \phi](k_E - \underline{k}_E)$$

or, using (5), this is

$$\omega = R_F P + \phi(\tilde{k}_E - \underline{k}_E) + [(1-\tau_K)(r_E - r_F) - \phi](k_E - \underline{k}_E),$$

which is the same expression for ω as in the equilibrium financial contract.

For the two contracts to be equivalent, the level of debt $\tilde{b}$ issued by the entrepreneur in the equity and debt contract is uniquely determined by (A.10) and (A.11), which imply $\tilde{b} = b - \frac{sE_\xi[\omega]}{(1-s)R_F}$. It remains to show that $R_F\tilde{b}$ is less than or equal to the value of the projects for the worst possible realization of ξ. That is, for all ξ it holds that $V_P - R_F\tilde{b} \geq 0$. Equation (A.11) implies that this holds as long as $\omega \geq 0$ for all ξ, which must be true since an entrepreneur's consumption is proportional to ω and consumption is non-negative. □

B Continuous Time Model

B.1 Environment and Equilibrium with Period Length Δ

Let $\Delta \in (0, 1]$ denote the length of a period. All assumptions are as in the main text except where specified here. Over a period of length Δ, agents discount future consumption at rate $(1-\rho\Delta)$, die with probability $\gamma\Delta$, capital depreciates at rate $\delta\Delta$, entrepreneurs draw a new productivity θ with probability $\lambda_\theta\Delta$ and entrepreneurs' idiosyncratic shocks entail that $\tilde{k}_{i,E,t} = q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t})k_{i,E,t}$, where ξ is drawn from H_ξ and the function q satisfies

$$q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t}) = \begin{cases} k_{E,i,t} & \text{if } k_{E,i,t} \leq \underline{k}_E \\ k_{E,i,t} + (1-\underline{\epsilon})\left(\exp\left(\frac{\varphi\xi_{i,t}\sqrt{\Delta}}{\sqrt{\theta_{i,t}}} - \frac{\varphi^2\Delta}{2\theta_{i,t}}\right) - 1\right)(k_{E,i,t} - \underline{k}_E) & \text{if } k_{E,i,t} > \underline{k}_E \end{cases} \tag{B.1}$$

Risky projects produce $\tilde{k}_E\Delta$ units and risk-free projects produce $k_F\Delta$ units. The government sets taxes τ_N, τ_K and $\tau_W\Delta$ per period, and has to finance exogenous expenditure $\overline{G}\Delta$.

An entrepreneur's expected lifetime utility is given by

$$\mathbb{E}_t\left[\sum_{j=0}^{\infty}(1-\rho\Delta)^j(1-\gamma\Delta)^j\log\left(c_{i,t+j\Delta}\right)\Delta + z_i\Delta\right]$$

and $P_{i,t}$ is

$$P_{i,t} = a_{i,t} + \sum_{j=0}^{\infty}\left[\frac{[(r_{E,t}-\delta)(1-\tau_{K,t+j}) - \tau_{W,t} - \tilde{R}_{F,t}]\underline{k}_E\Delta(1-\tau_{K,t+j})(1-\gamma\Delta)^j}{\prod_{k=0}^{j}1+\tilde{R}_{F,t+k}\Delta}\right] = a_{i,t} + F_t. \tag{B.2}$$

A worker's preferences are described by the lifetime utility function

$$\sum_{j=0}^{\infty}(1-\rho\Delta)^j(1-\gamma\Delta)^j\log\left(c^N_{t+j\Delta}\right)\Delta.$$

The worker's problem for a given period length Δ is

$$V^N(a^N, X) = \max\left(\log\left(c^N\right)\Delta + (1-\rho\Delta)(1-\gamma\Delta)V^N\left(a^{N\prime}, X'\right)\right) \tag{B.3}$$

subject to $c^N\Delta + (1-\gamma\Delta)a^{N\prime} = w\Delta(1-\tau_N) + (1+(R_F-1)\Delta)a^N$ and non-negativity constraints on c^N, $a^{N\prime}$.

The entrepreneur's problem is to solve

$$V(P,\theta,X) = \sup\int_\xi\left(\log(c)\Delta + (1-\rho\Delta)(1-\gamma\Delta)E\left[V(P',\theta',X')\middle|\theta\right]\right)\mathrm{d}H_\xi(\xi),$$

subject to

$$\begin{aligned}\omega &= [\phi - 1 - (1-\tau_K)r_E\Delta + \delta\Delta]k_H - (\hat{b} - (1+\tilde{R}_F\Delta)b) + (1+\tilde{R}_F\Delta)P \\ &+ [1+(1-\tau_K)\, r_E\Delta - \delta\Delta](\tilde{k}_E - \underline{k}_E) + \left(\tau_K\delta\Delta - \tau_W\Delta - 1 - \tilde{R}_F\Delta\right)(k_E - \underline{k}_E) \\ &+ [-\tilde{R}_F + (1-\tau_K)\,(r_F - \delta) - \tau_W]\Delta k_F,\end{aligned}$$

the incentive compatibility constraint

$$((1-\tau_K)\, r_E\Delta + (1-\delta\Delta))\,\frac{\partial\tilde{k}_E(P,\theta,\xi,X)}{\partial\xi} \geq \phi\frac{\partial\tilde{k}_E(P,\theta,\xi,X)}{\partial\xi} + \frac{\partial\hat{b}(P,\theta,\xi,X)}{\partial\xi},$$

the break-even condition for the banks

$$\int_\xi \hat{b}\,(a,\theta,\epsilon,X_t)\,\mathrm{d}H_\xi\,(\xi) = (1+\tilde{R}_F\Delta)b\,(P,\theta,X_t)\,,$$

and non-negativity constraints on k_E, k_F, k_H, c, ω and P'.

The fraction of newborn agents who choose to become workers each period is given by the probability $H_z(z \leq z^*)$, where z^* is the cutoff z that satisfies

$$E_\theta[V(F_t,\theta,X_t)] + \frac{z^*\Delta}{1-(1-\rho\Delta)(1-\gamma\Delta)} = V^N(F_t^N, X_t). \tag{B.4}$$

The aggregate number of workers evolves according to the number of workers who die and the number of newborn agents who become workers

$$N_{t+1} = (1-\gamma\Delta)N_t + \gamma\Delta H_z(z^*). \tag{B.5}$$

B.2 Solution to the Worker's and Entrepreneur's Problem with Period Length Δ

Following the derivation from the discrete time case, where $\Delta = 1$, the solution to the worker's problem is given by

$$c^N\Delta = (\gamma+\rho-\gamma\rho\Delta)\left(1+\tilde{R}_F\Delta\right)\Delta P^N, \tag{B.6}$$

$$P^{N'} = (1-\rho\Delta)\,(1+\tilde{R}_F\Delta)P^N, \tag{B.7}$$

where $P_{i,t}^N := a_{i,t}^N + \underbrace{\sum_{j=0}^{\infty}\left[\frac{w_{t+j}\Delta(1-\tau_{N,t+j})(1-\gamma\Delta)^j}{\prod_{k=0}^{j} 1+\tilde{R}_{F,t+k}\Delta}\right]}_{=:F_t}.$

The solution to the entrepreneur's between period problem is given by

$$c\Delta = (1-(1-\rho\Delta)(1-\gamma\Delta))\omega = (\gamma+\rho+\gamma\rho\Delta)\,\Delta\omega, \tag{B.8}$$

$$P' = (1-\rho\Delta)\omega = \frac{1}{1-\gamma\Delta}(\omega - c\Delta). \tag{B.9}$$

The entrepreneur's within period problem is to choose $k_E\,(P,\theta,X)$ and $\omega\,(P,\theta,X)$ to solve

$$\sup \int_\xi \log\left(\underline{\omega} + \phi(\tilde{k}_E - \underline{k}_E)\right) \mathrm{d}H_\xi(\xi), \tag{B.10}$$

subject to the constraints

$$\underline{\omega} = (-\phi + (r_E - r_F)(1-\tau_K)\Delta)\,(k_E - \underline{k}_E) + (1+\tilde{R}_F\Delta)P \tag{B.11}$$

$$k_E \geq \underline{k}_E \tag{B.12}$$

$$0 \leq \underline{\omega} + \phi\underline{\epsilon}(k_E - \underline{k}_E), \tag{B.13}$$

and where $\tilde{k}_E = q(\theta,\xi,k_E)$. The following proposition summarizes the optimal decisions.

Proposition 2. *In equilibrium, the entrepreneur's problem has a unique solution for $c(P,\theta,\epsilon,X)$, $P'(P,\theta,\epsilon,X)$, $\omega(P,\theta,\epsilon,X)$ and $k_E(P,\theta,X)$ which depends continuously on the parameters. The entrepreneur's optimal choice of k_E is*

$$k_E(P,\theta,X) - \underline{k}_E = \frac{S_\theta^{-1}\left(\max\left\{0;\min\left\{\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi};S_\theta^\star\right\}\right\}\right)P\left(1+\tilde{R}_F\Delta\right)}{\phi - (r_E-r_F)\Delta(1-\tau_K)S_\theta^{-1}\left(\max\left\{0;\min\left\{\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi};S_\theta^\star\right\}\right\}\right)}, \tag{B.14}$$

where $S_\theta^\star = S_\theta\left(\frac{1}{1-\underline{\epsilon}}\right)$. For any equilibrium values of r_E, r_F, the entrepreneur's choices entail

$$\omega = (\phi\,(\epsilon-1) + (1-\tau_K)(r_E-r_F)\Delta)\,(k_E - \underline{k}_E) + \left(1+\tilde{R}_F\Delta\right)P \tag{B.15}$$

$$c = (\gamma+\rho+\gamma\rho\Delta)\omega \tag{B.16}$$

$$P' = (1-\rho\Delta)\omega, \tag{B.17}$$

where

$$\epsilon = 1 + (1-\underline{\epsilon})\left(\exp\left(\frac{\varphi\xi_{i,t}\sqrt{\Delta}}{\sqrt{\theta_{i,t}}} - \frac{\varphi^2\Delta}{2\theta_{i,t}}\right) - 1\right), \tag{B.18}$$

$$S_\theta(x) = 1 - \frac{\int_\xi \left(1+x(\epsilon-1)\right)^{-1} \epsilon H_\xi\,(\xi)\,d\epsilon}{\int_\xi \left(1+x(\epsilon-1)\right)^{-1} H_\xi\,(\xi)\,d\epsilon}, \quad \forall x \in \left[0, \frac{1}{1-\underline{\epsilon}}\right], \tag{B.19}$$

and

$$S_\theta^\star = S_\theta\left(\frac{1}{1-\underline{\epsilon}}\right). \tag{B.20}$$

Here $S_\theta : \left[0, \frac{1}{1-\epsilon}\right] \to [0, S_\theta^\star]$ *is a differentiable and strictly increasing function.*

Proof. Note first that the definition of ϵ implies that $\tilde{k}_E - \underline{k}_E = \epsilon(k_E - \underline{k}_E)$. Then, the derivative of the entrepreneur's objective function with respect to k_E is

$$\frac{\partial}{\partial k_E}\int_\xi \log\left(\underline{\omega} + \phi\epsilon(k_E - \underline{k}_E)\right) \mathrm{d}H_\xi(\xi) = \frac{\int_\xi \left(1 - x + x\epsilon\right)^{-1}\left(\phi(\epsilon-1) + (r_E - r_F)\Delta\right)\mathrm{d}H_\xi(\xi)}{\underline{\omega} + \phi(k_E - \underline{k}_E)},$$

where

$$x = \frac{\phi(k_E - \underline{k}_E)}{\underline{\omega} + \phi(k_E - \underline{k}_E)}. \tag{B.21}$$

Here we used that $\underline{\omega} + \phi(k_E - \underline{k}_E) > 0$ at any feasible k_E. Since $k_E - \underline{k}_E \geq 0$, equation (B.21) in turn implies that $x \geq 0$ at any feasible choice. Furthermore, $x \leq \frac{1}{1-\underline{\epsilon}}$, or else the entrepreneur's consumption could be negative.

Now we show that x is monotonically increasing in k_E. Using the definitions of $\underline{\omega}$ and x in equations (B.13) and (B.21), we obtain that $\frac{\partial x}{\partial k_E} \propto \left(1 + \tilde{R}_F\Delta\right) P$. Since the entrepreneur can convert capital into consumption at rate ϕ and the risk-free return is $(1+\tilde{R}_F\Delta)$, entrepreneurs put a positive amount of capital in the risk-free sector only if $(1 + \tilde{R}_F\Delta) > 0$. Moreover $P > 0$, so $\frac{\partial x}{\partial k_E} > 0$. So x is monotonically increasing in k_E, and $x = 0$ when $(k_E - \underline{k}_E) = 0$, while $x = \frac{1}{1-\underline{\epsilon}}$ is the highest possible k_E that guarantees non-negative consumption.

Using that $x \in [0, \frac{1}{1-\underline{\epsilon}}]$, the derivative of the objective function can be rearranged to

$$\frac{\phi}{\underline{\omega} + \phi(k_E - \underline{k}_E)}\int_\xi \left(1 + x(\epsilon - 1)\right)^{-1}\mathrm{d}H_\xi(\xi)\left(\frac{(r_E - r_F)\Delta(1-\tau_K)}{\phi} - S_\theta(x)\right),$$

where

$$S_\theta(x) = 1 - \frac{\int_\xi \left(1 + x(\epsilon - 1)\right)^{-1}\epsilon \mathrm{d}H_\xi\left(\xi\right)}{\int_\xi \left(1 + x(\epsilon - 1)\right)^{-1}\mathrm{d}H_\xi\left(\xi\right)}.$$

Since $\underline{\omega} + \phi(k_E - \underline{k}_E) > 0$ and $x^{-1} + \epsilon - 1 \geq 0$ for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, with strict inequality for $\epsilon > \underline{\epsilon}$, it follows that the sign of this derivative is given by the sign of

$$\frac{(r_E - r_F)\Delta(1-\tau_K)}{\phi} - S_\theta(x), \tag{B.22}$$

expression that is equal to 0 at the optimal interior k_E. It can then be shown that $S_\theta(x)$ is strictly increasing for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$. Equation (B.19) immediately implies that S is continuous over $x \in [0, \frac{1}{1-\underline{\epsilon}}]$ and so $S_\theta(x)$ is also strictly increasing over $x \in [0, \frac{1}{1-\underline{\epsilon}}]$.

It then follows that there are three cases. If $\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi} \leq S_\theta(0)$, then the entrepreneur optimally chooses the corner solution $k_E - \underline{k}_E = x = 0$. If $\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi} \geq S_\theta\left(\frac{1}{1-\underline{\epsilon}}\right)$, then the entrepreneur optimally chooses the corner solution $x = \frac{1}{1-\underline{\epsilon}}$, which corresponds to the highest possible choice of k_E. If $S_\theta(0) < \frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi} < S_\theta\left(\frac{1}{1-\underline{\epsilon}}\right)$, then, by the intermediate value theorem, there is a unique x satisfying the first order condition with respect to k_E. We can group these three cases as follows

$$x = \begin{cases} S_\theta^{-1}(0) & \text{if } \frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi} \leq 0 \\ S_\theta^{-1}\left(\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi}\right) & \text{if } \frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi} \in (0, S_\theta^\star) \\ S_\theta^{-1}(S_\theta^\star) & \text{if } \frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi} \geq S_\theta^\star \end{cases}$$

Combining this with (B.21) to solve for k_E, and simplifying, we have

$$k_E - \underline{k}_E = \frac{S_\theta^{-1}\left(\max\left\{0; \min\left\{\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi}; S_\theta^\star\right\}\right\}\right) P\left(1+\tilde{R}_F\Delta\right)}{\phi - (r_E-r_F)\Delta(1-\tau_K) S_\theta^{-1}\left(\max\left\{0; \min\left\{\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi}; S_\theta^\star\right\}\right\}\right)}$$

From equations (4) and (B.11), we have that

$$\omega = \left(\phi\left(\epsilon - 1\right) + (1-\tau_K)(r_E - r_F)\Delta\right)\left(k_E - \underline{k}_E\right) + \left(1 + \tilde{R}_F\Delta\right) P.$$

Combining with equations (B.8) and (B.9) yields all the results of the proposition. □

B.3 Optimal Choices of Workers and Entrepreneurs

Taking the limit when $\Delta \to 0$, yields the following optimal decision rules.

Proposition 3. *In equilibrium, the unique solution of the worker's problem is*

$$c^N = (\rho+\gamma)P^N \qquad \text{and} \qquad dP^N = \left\{\left[\tilde{R}_F + \gamma\right] P^N - c^N\right\} dt,$$

where $\tilde{R}_F = R_F - 1$ denotes the net risk-free rate of return. If $r_E > r_F$ then the unique equilibrium solution of the entrepreneur's problem is

$$\begin{aligned} k_H &= 0, \qquad k_E = \underline{k}_E + P\hat{k}_E(\theta), \qquad c = (\rho+\gamma)P \\ dP &= \left[\left(\tilde{R}_F + \gamma\right) P + \left(k_E - \underline{k}_E\right)\left(r_E - r_F\right)\left(1-\tau_K\right) - c\right] dt + \frac{(k_E - \underline{k}_E)\phi\left(1-\underline{\epsilon}\right)\varphi}{\sqrt{\theta}} dW, \end{aligned}$$

where dW is the difference of a standard Brownian motion and where

$$\hat{k}_E(\theta) \equiv \frac{1}{\phi(1-\underline{\epsilon})} \times \min\left[\frac{(r_E - r_F)(1-\tau_K)\theta}{\phi(1-\underline{\epsilon})\varphi^2}; 1\right]. \tag{B.23}$$

Proof. The proof of this proposition makes use of the following two lemmas. We omit the proofs of the lemmas, but note that they are avaiable upon request.

Lemma 2. *The following holds, for any $x \in \left[0; \frac{1}{1-\underline{\epsilon}}\right]$: $\lim_{\Delta\to 0} \frac{S_\theta(x)}{\Delta} = \frac{(1-\underline{\epsilon})^2\varphi^2 x}{\theta}$.*

Lemma 3. *For any $z \in (-\infty, \infty)$, it holds that*

$$\lim_{\Delta\to 0} S_\theta^{-1}(\max\{0; \min\{z\Delta; S_\theta^\star\}\}) = \max\left\{0; \min\left\{\frac{z\theta}{(1-\underline{\epsilon})^2\varphi^2}; \frac{1}{1-\underline{\epsilon}}\right\}\right\}. \tag{B.24}$$

Turning to the proof of the proposition, we first derive the worker's continuous time solution. Taking the limit of equation (B.6) as $\Delta \to 0$, we obtain c^N. Now consider $\frac{P^{N'} - P^N}{\Delta} = \left[(1-\rho\Delta)\tilde{R}_F - \rho\right] P^N$. Taking the limit of this as $\Delta \to 0$, we get $dP^N = \left\{\left[\tilde{R}_F - \rho\right] P^N\right\} \mathrm{d}t$. Using the solution for c^N, we obtain dP^N.

Now we derive the entrepreneur's continuous time solution. Taking the limit of equation (B.14) as $\Delta \to 0$, and using Lemma 3, implies (B.23). Combining (B.15) and (B.16) implies

$$c = (\rho + \gamma - \rho\gamma\Delta)\left[P\left(1 + \tilde{R}_F\Delta\right) + (k_E - \underline{k}_E)\left(\phi\left(\epsilon - 1\right) + (1-\tau_K)(r_E - r_F)\Delta\right)\right].$$

Taking the limit of this as $\Delta \to 0$ and noting that, as $\Delta \to 0$, $\epsilon \to 1$ in probability, we obtain c. Finally, we note that Proposition 2 implies that when Δ is close to 0

$$P' - P \;=\; \left[\left(\tilde{R}_F + \gamma\right) P + (k_E - \underline{k}_E)\left(r_E - r_F\right)\left(1 - \tau_K\right) - c\right]\Delta + \frac{(k_E - \underline{k}_E)\phi\left(1 - \underline{\epsilon}\right)\varphi}{\sqrt{\theta}}\xi\sqrt{\Delta}$$

Since ξ has mean 0 and variance 1, it follows that the variance of $P' - P$ is proportional to Δ. Standard arguments then imply that, as $\Delta \to 0$, P evolves according to an Ito process. Replacing Δ with $\mathrm{d}t$ and $\xi\sqrt{\Delta}$ with $\mathrm{d}W$, in the expression above, we obtain dP.

Finally, all entrepreneurs choose $k_H = 0$. The entrepreneur will not hide capital if $\phi < (1-\tau_K)r_E + (1-\delta)$. With period length Δ, this inequality becomes $\phi < (1-\tau_K)r_E\Delta + 1 - \delta\Delta$, In the limit as $\Delta \to 0$, this is $\phi < 1$, which is satisfied since $\phi \in (0, 1)$.

□

B.4 Characterization of the Steady State Equilibrium

We define the following post-tax prices: $\tilde{R}_F = R_F - 1$, the post-tax wage $\tilde{w} \equiv w(1 - \tau_N)$ and post-tax excess return to the risky project $\tilde{r}_X \equiv (r_E - r_F)(1 - \tau_K)$.

The steady state equilibrium of the economy can be characterized in terms of aggregate variables aggregate variables $\{Y^*, K^*, K_E^*, C^*, N^*\}$ and post-tax prices $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ according to Proposition 4 below. In the interest of space, we omit the proof of the proposition, which is available upon request. We simply note here that to prove the proposition we first characterize a steady state equilibrium in terms of aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, pre-tax prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_N^*\}$ and show the equivalence between the two characterizations of the steady state.

Proposition 4. *There exists a steady state which is consistent with the values of aggregate variables* $\{Y^*, K^*, K_E^*, C^*, N^*\}$*, functions* $\mu(\theta)$*,* $\hat{k}_E(\theta)$ *and post-tax prices* $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*\right\}$*, iff*

$$C^* = \frac{\rho+\gamma}{\tilde{R}_F^*+\gamma}\left(Y^* - \delta K^* - \overline{G} - \tilde{r}_X^*(K_E^* - (1-N^*)\underline{k}_E) + \gamma K^*\right) \tag{B.25}$$

$$C^* = Y^* - \delta K^* - \overline{G} = N^*\tilde{w}^* + \tilde{r}_X^* K_E^* + \tilde{R}_F^* K^* \tag{B.26}$$

$$N^* = H_z\left(\tilde{w}^* - \log \tilde{r}_x^*\underline{k}_E - \frac{\mathbb{E}[\theta]}{\rho+\gamma}\left(\tilde{r}_X^*\hat{k}(1) - \frac{(\phi(1-\underline{\epsilon})\hat{k}(1)\varphi)^2}{2}\right)\right) \tag{B.27}$$

$$K_E^* = \underline{k}_E(1-N^*)\left(1 + \frac{\frac{\gamma}{\gamma+\tilde{R}_F^*}\tilde{r}_X^*\int \hat{k}_E(\theta)\mu(\theta)\,d\theta}{\rho+\gamma-\tilde{R}_F^* - \tilde{r}_X^*\int \hat{k}_E(\theta)\mu(\theta)\,d\theta}\right) \tag{B.28}$$

$$\mu(\theta) = \frac{h_\theta(\theta)}{1 - \frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}}\left(\int_0^1 \frac{h_\theta(\theta)}{1 - \frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}}\,d\theta\right)^{-1} \tag{B.29}$$

$$\hat{k}_E(\theta) = \frac{\tilde{r}_X^*\theta}{(\phi(1-\underline{\epsilon})\varphi)^2} = \hat{k}_E(1)\theta \tag{B.30}$$

and $Y^* = f\left(K_E^*, K^* - K_E^*, N^*\right)$*,* $\underline{k}_E(1-N^*) < K_E^* < K^*$*,* $\tilde{R}_F + \gamma > 0$*,* $\lambda_\theta + \rho + \gamma - \tilde{R}_F^* > \tilde{r}_X^*\hat{k}_E(1) > 0$*.*

B.5 Characterizing Partial Equilibrium Elasticities of Tax Changes

We define and compute the partial equilibrium derivative $\frac{\partial X}{\partial \tau_j}$ for each $X \in \{K; K_E; N; Y\}$ and $j \in \{K; W\}$. To fix ideas, we first discuss how to compute $\frac{\partial Y}{\partial \tau_K}$, that is the marginal effect of τ_K on the steady state value of Y, holding fixed pre-tax prices. All partial equilibrium elasticities can then be defined in a similar fashion. We write all steady state variables without asterisks, for simplicity.

To define $\frac{\partial Y}{\partial \tau_K}$, first differentiate $\tilde{r}_X$ and $\tilde{R}_F$ with respect to τ_K, holding constant pre-tax prices to obtain $\frac{\partial \tilde{r}_X}{\partial \tau_K} = r_E - r_F$ and $\frac{\partial \tilde{R}_F}{\partial \tau_K} = r_F - \delta$. Then, we make use of Proposition 4 and invoke the implicit function theorem to show that in the neighborhood of some initial steady state $\mathscr{S}$, we can write the steady state values of seven of the equilibrium variables as continuously differentiable functions of the two variables $\tilde{r}_X$ and $\tilde{R}_F$.[30] Therefore, we can treat K, Y, K_E and C as functions of $\tilde{r}_X$ and $\tilde{R}_F$, and use the equations of Proposition 4 to compute the partial derivatives $\frac{\partial Y}{\partial \tilde{r}_X}$ and $\frac{\partial Y}{\partial \tilde{R}_F}$. Combining these partial derivatives with the values of $\frac{\partial \tilde{r}_X}{\partial \tau_K}$ and $\frac{\partial \tilde{R}_F}{\partial \tau_K}$, we can write

$$\frac{\partial Y}{\partial \tau_K} = -(r_E - r_F)\frac{\partial Y}{\partial \tilde{r}_X} - (r_F - \delta)\frac{\partial Y}{\partial \tilde{R}_F},$$

and so $e^Y_{\tau_K} = -\frac{1-\tau_K}{Y}\left((r_E - r_F)\frac{\partial Y}{\partial \tilde{r}_X} + (r_F - \delta)\frac{\partial Y}{\partial \tilde{R}_F}\right)$.

The same logic can be used to define $\frac{\partial X}{\tau_j}$ for any aggregate variable X, and for $j \in \{K; W\}$

$$e^X_{\tau_K} = \frac{1-\tau_K}{X}\frac{\partial X}{\partial \tau_K} = -\frac{\tilde{r}_X}{X}\frac{\partial X}{\partial \tilde{r}_X} - \frac{(1-\tau_K)(r_F - \delta)}{X}\frac{\partial X}{\partial \tilde{R}_F} \tag{B.31}$$

$$e^X_{\tau_W} = \frac{1}{X}\frac{\partial X}{\partial \tau_W} = -\frac{1}{X}\frac{\partial X}{\partial \tilde{R}_F}, \tag{B.32}$$

and the derivatives $\frac{\partial X}{\partial \tilde{r}_X}$ and $\frac{\partial X}{\partial \tilde{R}_F}$ can all be defined as described above.

B.5.1 Output Elasticity

Proposition 5. *The partial equilibrium elasticity of steady state output Y with respect to the tax rates τ_j, $j \in \{K, W\}$, adjusting τ_N to balance the government's budget is*

$$e^Y_{\tau_j} = \overbrace{\left((r_E - r_F)\frac{K_E}{Y} + \frac{r_F K}{Y}\right)}^{\text{Capital's share}} e^K_{\tau_j} + \frac{wN}{Y}e^N_{\tau_j} + \overbrace{(r_E - r_F)\frac{K_E}{Y}(e^{K_E}_{\tau_j} - e^K_{\tau_j})}^{\text{Reallocation Effect}}.$$

Proof. Differentiate $Y = f(K_E, K, N)$ with respect to $\tilde{r}_X$ and $\tilde{R}_F$, using that the marginal product of inputs is equal to factor prices. Thus, for $x \in \{\tilde{r}_X; \tilde{R}_F\}$

$$\frac{\partial Y}{\partial x} = (r_E - r_F)\frac{\partial K_E}{\partial x} + r_F\frac{\partial K}{\partial x} + w\frac{\partial N}{\partial x}$$

Substitute $\frac{\partial Y}{\partial \tilde{r}_X}$ and $\frac{\partial Y}{\partial \tilde{R}_F}$ into equations (B.31) and (B.32). The result follows immediately. □

[30]This holds if the relevant Jacobian is invertible, which is the case outside of knife-edge situations.

B.5.2 Risky Capital Elasticity

Proposition 6. *The partial equilibrium elasticity of K_E with respect to the tax rates τ_j, $j \in \{K, W\}$, adjusting τ_N to balance the government's budget is*

$$e^{K_E}_{\tau_j} = \left(1 - \frac{\underline{k}_E(1-N)}{K_E}\right) M_{K_E}(e^{\hat{k}_E}_{\tau_j} + e^{\mathbb{P}}_{\tau_j}) - \frac{N e^N_{\tau_j}}{1-N},$$

where

$$\begin{aligned} M_{K_E} &= \tfrac{K_E}{\underline{k}_E(1-N)}\left(1 + \tfrac{\tilde{R}_F}{\gamma}\left(1 - \tfrac{\underline{k}_E(1-N)}{K_E}\right)\right) & > 1, \\ e^{\mathbb{P}}_{\tau_W} &= -\tfrac{1}{M_{KE}(\gamma+\tilde{R}_F)} - \tfrac{1}{\rho+\gamma-\tilde{R}_F} & < 0 \\ e^{\mathbb{P}}_{\tau_K} &= -1 + (1-\tau_K)(r_F - \delta)e^{\mathbb{P}}_{\tau_W} & < 0, \end{aligned}$$

and $e^{\hat{k}_E}_{\tau_K} = (1-\tau_K)\frac{\partial}{\partial \tau_K}\log\left(\int_\theta \mu(\theta)\hat{k}_E(\theta)\,d\theta\right)$, with an analogous definition of $e^{\hat{k}_E}_{\tau_W}$.

Proof. Differentiate (B.28) with respect to $\tilde{r}_X$ and $\tilde{R}_F$. Substitute into equations (B.31) and (B.32) and rearrange. The result follows. □

We next establish that for sufficiently large λ_θ, $e^{\hat{k}_E}_{\tau_K}$ and $e^{\hat{k}_E}_{\tau_W}$ are -1 and 0, respectively.

Lemma 4. *In the limit as λ_θ approaches infinity, the elasticities $e^{\hat{k}_E}_{\tau_K}$ and $e^{\hat{k}_E}_{\tau_W}$ satisfy $e^{\hat{k}_E}_{\tau_K} \to -1$ and $e^{\hat{k}_E}_{\tau_W} \to 0$.*

Proof. Using (B.30), since $\tilde{r}_X = (r_E - r_F)(1-\tau_K)$, it follows that

$$(1-\tau_K)\frac{\partial \log \hat{k}_E(1)}{\partial \tau_K} = -1 \quad \text{and} \quad \frac{\partial \log \hat{k}_E(1)}{\partial \tau_W} = 0.$$

Then, to prove the Lemma, it is sufficient to show that, for $j \in \{K; W\}$,

$$\lim_{\lambda_\theta \to \infty} \frac{\partial}{\partial \tau_j}\left(\hat{k}_E(1)^{-1}\int_\theta \mu(\theta)\hat{k}_E(\theta)\mathrm{d}\theta\right) = 0,$$

which follows from $\hat{k}_E(\theta) = \theta\hat{k}_E(1)$ if, for all $\theta \in [0,1]$, $\lim_{\lambda_\theta\to\infty}\frac{\partial\mu(\theta)}{\partial\tau_j} = 0$. Using equation (B.29), this follows from the quotient rule of differentiation if, for all $\theta \in [0,1]$,

$$\lim_{\lambda_\theta \to \infty} \frac{\partial}{\partial \tau_j}\left(\frac{h_\theta(\theta)}{1 - \left(\frac{\tilde{r}^*_X \hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}^*_F}\right)}\right) = 0,$$

since, in that case, the derivative of the integral over θ of $\frac{h_\theta(\theta)}{1-\left(\frac{\tilde{r}^*_X \hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}^*_F}\right)}$ must also approach

0 in the limit. Now,

$$\frac{\partial}{\partial \tau_j}\left(\frac{h_\theta(\theta)}{1-\left(\frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}\right)}\right)=\frac{h_\theta(\theta)\theta}{\left(1-\left(\frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}\right)\right)^2}\frac{\partial}{\partial \tau_j}\left(\frac{\tilde{r}_X^*\hat{k}_E(1)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}\right)$$

Using (B.30), it follows that

$$\lim_{\lambda_\theta\to\infty}\frac{h_\theta(\theta)\theta}{\left(1-\left(\frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}\right)\right)^2}=h_\theta\theta \quad \text{and} \quad \lim_{\lambda_\theta\to\infty}\frac{\partial}{\partial \tau_j}\left(\frac{\tilde{r}_X^*\hat{k}_E(1)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}\right)=0.$$

□

B.5.3 Capital Elasticity

Proposition 7. *The partial equilibrium elasticity of steady state capital stock K with respect to the tax rates τ_j, $j \in \{K, W\}$, adjusting τ_N to balance the government's budget is*

$$e_{\tau_j}^K = \frac{\frac{K_E}{K}e_{\tau_j}^{K_E}\left(r_E-r_F\right)\left(1-\tau_K MPC\right)+e_{\tau_j}^{S_{SUB}}+e_{\tau_j}^N\left[w(1-MPC)+\tilde{r}_X\underline{k}_E MPC\right]\frac{N}{K}}{\gamma MPC-\left(r_F-\delta\right)\left(1-MPC\right)},$$

where $e_{\tau_K}^{S_{SUB}}=-\left(\frac{\tilde{r}_X(K_E-(1-N)\underline{k}_E)}{K}\right)MPC-\frac{C}{K}e_{\tau_K}^{MPC}$ *and* $e_{\tau_W}^{S_{SUB}}=-\frac{C}{K}e_{\tau_W}^{MPC}$.

Proof. Let $x \in \{\tilde{r}_X; \tilde{R}_F\}$. Combining equations (B.25) and (B.26) and differentiating, we obtain

$$\frac{\partial Y}{\partial x}-\delta\frac{\partial K}{\partial x}=MPC\frac{\partial}{\partial x}\left(Y-\delta K-\overline{G}-\tilde{r}_X(K_E-(1-N)\underline{k}_E)+\gamma K\right)+\frac{C}{MPC}\frac{\partial MPC}{\partial x},$$

where $MPC=\frac{\rho+\gamma}{\tilde{R}_F+\gamma}$ From the proof of Proposition 5 above, we have

$$\frac{\partial Y}{\partial x}=(r_E-r_F)\frac{\partial K_E}{\partial x}+r_F\frac{\partial K}{\partial x}+w\frac{\partial N}{\partial x}$$

Combining these expressions, substituting into equations (B.31) and (B.32) rearranging, we obtain the desired result. □

B.5.4 Labor Elasticity

We now derive the effect of a change in taxes on N. The algebraic derivation is similar to the one we use in deriving the optimal taxes. Steady-state N depends on $\tilde{w}$, $\tilde{r}_X$ and $\hat{k}_E$, which itself depends on $\tilde{r}_X$. Then a change in τ_W only affects N through its effect on $\tilde{w}$, as it induces a budget balancing change in τ_N. Differentiating equation (B.27) with respect to

τ_W, we obtain that

$$\frac{\partial N}{\partial \tau_W} = \frac{H_z'(z^\star)}{\tilde{w}}\frac{\partial \tilde{w}}{\partial \tau_W} = -\frac{H_z'(z^*)}{N}\frac{N}{1-\tau_N}\frac{\partial \tau_N}{\partial \tau_W} = \frac{-e_{\tilde{w}}^N N}{1-\tau_N}\frac{\partial \tau_N}{\partial \tau_W}, \tag{B.33}$$

where we used that $z^\star$ is the argument of the cdf $H_z(\cdot)$ in equation (B.27) and where $e_{\tilde{w}}^N = \frac{H'(z^*)}{N}$ is the partial equilibrium elasticity of N with respect to $\tilde{w}$.

We then use the government's budget constraint to infer $\frac{\partial \tau_N}{\partial \tau_W}$, which can be rewritten as

$$\overline{G} = \tau_N B_{\tau_N} + \tau_K B_{\tau_K} + \tau_W B_{\tau_W},$$

where B_{τ_j} is the tax base for the tax τ_j, so that $B_{\tau_N} = wN$, $B_{\tau_K} = (r_E - r_F)K_E + (r_F - \delta)K$ and $B_{\tau_W} = K$. Differentiating with respect to τ_W and rearranging, we obtain that

$$-B_{\tau_N}\frac{\partial \tau_N}{\partial \tau_W} = B_{\tau_W} + \sum_{j\in\{K;W;N\}} \tau_j \frac{\partial B_{\tau_j}}{\partial \tau_W}.$$

Using the definitions of the B_{τ_j} above, we can write $\frac{\partial B_{\tau_m}}{\partial \tau_j}$ as a function of the elasticities of K_E, K and N with respect to taxes. For instance

$$\begin{aligned} e_{\tau_W}^{B_{\tau_N}} = \frac{1}{B_{\tau_N}}\frac{\partial B_{\tau_N}}{\partial \tau_W} &= e_{\tau_W}^N, \\ e_{\tau_W}^{B_{\tau_K}} = \frac{1}{B_{\tau_K}}\frac{\partial B_{\tau_K}}{\partial \tau_W} &= \left[1 - \frac{(r_F-\delta)K}{B_{\tau_K}}\right] e_{\tau_W}^{K_E} + \left[\frac{(r_F-\delta)K}{B_{\tau_K}}\right] e_{\tau_W}^K, \\ e_{\tau_W}^{B_{\tau_W}} = \frac{1}{B_{\tau_W}}\frac{\partial B_{\tau_W}}{\partial \tau_W} &= e_{\tau_W}^K. \end{aligned}$$

Substituting $\frac{\partial \tau_N}{\partial \tau_W}$ into equation (B.33) and using that $e_{\tau_W}^N = \frac{1}{N}\frac{\partial N}{\partial \tau_W}$, we obtain the partial equilibrium effect of τ_W on N, given below. The partial equilibrium effect of τ_K is similar, except that an additional term $e_{\tau_K}^{z^\star D}$ needs to be added, which represents that an increase in τ_K also raises N by directly decreasing the cutoff $z^\star$ for becoming an entrepreneur by lowering the relative expected lifetime income of entrepreneurs. Intuitively, τ_K falls relatively more on entrepreneurs than on workers because they can earn a return to capital greater than $\tilde{R}_F$. This discourages entry into entrepreneurship more than τ_W does.

Proposition 8. *The partial equilibrium elasticities of steady state aggregate labor N with respect to the tax rates τ_K and τ_W, assuming that τ_N adjusts to balance the government's* *budget are, respectively*

$$e^N_{\tau_W} = \left(\frac{1-\tau_N}{e^N_{\tilde{w}}} - \tau_N\right)^{-1} B^{-1}_{\tau_N}\left(B_{\tau_W} + \sum_{j\in\{K;W\}} \tau_j \frac{\partial B_{\tau_j}}{\partial \tau_W}\right)$$

$$e^N_{\tau_K} = \left(\frac{1-\tau_N}{e^N_{\tilde{w}}} - \tau_N\right)^{-1}\left((1-\tau_N)e^{z^\star D}_{\tau_K} + B^{-1}_{\tau_N}(1-\tau_K)\left[B_{\tau_W} + \sum_{j\in\{K;W\}} \tau_j \frac{\partial B_{\tau_j}}{\partial \tau_W}\right]\right),$$

where $e^{z^\star D}_{\tau_K} = 1 + \frac{\mathbb{E}[\theta]}{\rho+\gamma}(1-\tau_K)\frac{\partial}{\partial \tau_K}\left(\tilde{r}^*_X \hat{k}(1) - \frac{(\phi(1-\epsilon)\hat{k}(1)\varphi)^2}{2}\right)$.

B.6 Effects of Tax Changes on Welfare

The following lemma characterizes the effect of a marginal tax change on worker utility.

Lemma 5. *The change in worker steady state lifetime utility from a marginal change in taxes satisfies*

$$\tilde{w}\, dV^N(F^N, X) = \frac{1}{\rho+\gamma}\left(d\tilde{w} + \mathcal{A}^N d\tilde{R}_F\right),$$

where $\mathcal{A}^N = (\gamma + \tilde{R}_F)\int_{s=0}^{\infty} e^{-(\gamma+\tilde{R}_F)s} a^N_s ds$.

Proof. We focus first on the discrete time case. Since workers choose a^N_{s+1} optimally each period, envelope theorem arguments imply that we may calculate the resulting change in their lifetime utility as if workers continue to choose the same level of a^N_{s+1} each period irrespective of the tax change.[31] Then, the worker's budget constraint implies that the tax change has an effect on his welfare equivalent to increasing worker consumption by $\mathrm{d}c^N_s$ in each period s, where $\mathrm{d}c^N_s$ satisfies $\mathrm{d}c^N_s = \mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F a^N_s$, where $\mathrm{d}\tilde{w}$ and $\mathrm{d}\tilde{R}_F$ are the change in $\tilde{w}$ and $\tilde{R}_F$ as a result of the tax change. In such a case, the tax change increases the present value of the worker's lifetime resources by

$$\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s \mathrm{d}c^N_s = \left(\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s\right)\left(\mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \mathcal{A}^N\right)$$

where

$$\mathcal{A}^N = \frac{\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s a^N_s}{\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s}$$

[31] In particular, the total change in worker welfare is equal to the change holding a^N_{s+1} constant each period, plus the effect of the resulting changes in each period's choice of a^N_{s+1} on worker welfare. But the worker's first order condition implies that the latter effects must be zero.

is the average value of the worker's discounted lifetime assets. Applying envelope theorem arguments further, the change in worker lifetime utility from a small tax change must then be equivalent to the change in worker utility if the worker consumed all the extra resources $\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s \mathrm{d}c_s^N$ in the first period of their life, since on the margin, workers are indifferent about which period they consume each extra unit of lifetime resources they receive. That is, the change in welfare satisfies

$$\mathrm{d}V^N(F^N, X^*) = \frac{1}{c_0^N}\left(\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s\right)\left(\mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F\mathcal{A}^N\right),$$

where $\frac{1}{c_0^N}$ is the worker's marginal utility of consumption in the first year of her life.

Combining this with the definition of F^N and using that the worker consumes $c_0^N = [1-(1-\rho)(1-\gamma)]\left(1+\tilde{R}_F\right)F^N$, this simplifies to

$$\mathrm{d}V^N(F^N, X^*) = \frac{1}{\tilde{w}}\left(\frac{1}{[1-(1-\rho)(1-\gamma)]}\right)\left(\mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F\mathcal{A}^N\right).$$

Repeating the same arguments in the model with period length Δ and taking the limit as $\Delta \to 0$, we obtain the result of the lemma. □

Then, using the occupational choice condition $N = H_z(z^\star)$ and $e_{\tilde{w}}^N = \frac{H'(z^\star)}{N}$ we can express the effect of a marginal tax change on aggregate welfare as

$$\tilde{w}\mathrm{d}\mathcal{W} = \frac{1}{\rho+\gamma}\left(d\tilde{w} + \mathcal{A}^N\mathrm{d}\tilde{R}_F - \frac{(1-N)\tilde{w}\mathrm{d}N}{Ne_{\tilde{w}}^N}\right).$$

As is common in the literature, we focus on the percentage consumption equivalent welfare change, which we denote by Δ and which satisfies

$$\mathcal{W} + \mathrm{d}\mathcal{W} = \int_i\int_{s=0}^{\infty} e^{-(\rho+\gamma)}\log\left((1+\Delta)\,c_{i,t+s}\right)\mathrm{d}s\mathrm{d}i \equiv \mathcal{W} + \frac{\log(1+\Delta)}{\rho+\gamma}.$$

Thus, for small Δ, it follows that $\mathrm{d}\mathcal{W} = \frac{\Delta}{\rho+\gamma}$, and Δ satisfies $\tilde{w}\Delta = d\tilde{w} + \mathcal{A}^N\mathrm{d}\tilde{R}_F - \frac{(1-N)\tilde{w}\mathrm{d}N}{Ne_{\tilde{w}}^N}$.

We can study the partial equilibrium effects of tax changes on welfare using the same approach as in Section 4.1 and characterize the partial equilibrium effect of a small change in tax rate τ_j, $j \in \{K; W\}$ on welfare as

$$(1-\tau_N)wN\Delta = \left[B_{\tau_j} + \left(\sum_{m\in\{K;W;N\}}\tau_m\frac{\partial B_{\tau_m}}{\partial\tau_j}\right) - B_{\tau_j}^N N - \frac{(1-N)\,w\,(1-\tau_N)}{e_{\tilde{w}}^N}\frac{\partial N}{\partial\tau_j}\right]\mathrm{d}\tau_j,$$

where $B_{\tau_K}^N = (r_F - \delta)\mathcal{A}^N$ and $B_{\tau_W}^N = \mathcal{A}$.

B.7 Generality of the Optimal Tax Formula

The derivation of the optimal tax formula obtained in Proposition 1 does not depend on many specific features of the model, including the financial friction, the functional form determining entrepreneurial risk, and the logarithmic utility. The following proposition establishes the result.

Proposition 9. *Suppose that the model assumptions are as in Section 2 except that*

1. $q(\cdot,\cdot,\cdot)$ *is a continuously differentiable function, with* $\mathbb{E}_{\xi} q(\theta, \xi, k_E) = k_E$.
2. $\phi \geq 0$, *and there may be other financial friction which restricts entrepreneurs' choices.*
3. *The period utility function* $u_{i,t}$ *is given by*

$$u_{i,t}(c_{i,t}) = \begin{cases} \overline{u}(c^N_{i,t}) & \text{if } i \text{ is a worker} \\ \overline{u}(c^N_{i,t}) + z_{i,t} & \text{if } i \text{ is an entrepreneur} \end{cases}$$

where $\overline{u}$ *is a twice continuously differentiable function with* $\overline{u}'(\cdot) > 0$ *and* $\overline{u}''(\cdot) < 0$.

Suppose further that, as the period length approaches zero, no capital is hidden and there exists a steady state in which the value functions $V^N(a, X)$ *and* $V(a, \theta, X)$ *and all aggregate variables and elasticities referenced in Proposition 1 are finite and non-zero. Then, optimal taxes are given by the formula in Proposition 1.*

Proof. Repeating the steps in Appendix B.6, we obtain that the change in a newborn worker's lifetime utility from a change in taxes is given by

$$\mathrm{d}V^N(0, X) = u'(c^N_0) \left(\frac{F^N}{\tilde{w}}\right) \left[\mathrm{d}\tilde{w} + \mathcal{A}^N \mathrm{d}\tilde{R}_F\right], \tag{B.34}$$

where $\tilde{w}$, $\mathcal{A}^N$,F^N and $\tilde{R}_F$ are all defined as before. We omit reference to F^N in the value function and instead write it in terms of the newborn's assets, $a^N = 0$. We also omit asterisks denoting steady state variables. The only change from the formula obtain in Appendix B.6 is that $u'(c^N_0)$ is no longer the same as $\frac{1}{c^N_0}$.

As before, the condition for optimal occupational choice is

$$V^N(0, X) = \mathbb{E}_\theta V(0, \theta, X) + \frac{z^*}{\rho + \gamma}$$

This implies that the change in N from a change in taxes satisfies

$$\mathrm{d}N = H'_z(z^*)[\mathrm{d}V^N(0, X) - \mathrm{d}\mathbb{E}_\theta V_E(0, \theta, X)] \tag{B.35}$$

Since a change in $\tilde{w}$ does not directly affect V_E, this implies that the elasticity of N with respect to a change in $\tilde{w}$ satisfies

$$e^N_{\tilde{w}} = \frac{\tilde{w}}{N}\frac{\partial N}{\partial \tilde{w}} = \frac{H'_z(z^*)\tilde{w}}{N}\frac{\partial V^N}{\partial \tilde{w}}$$

Combining this with equation (B.34) and rearranging, we obtain

$$\mathrm{d}V^N(0,X) = \frac{N e^N_{\tilde{w}}}{H'_z(z^*)\tilde{w}}\left[\mathrm{d}\tilde{w} + \mathcal{A}^N \mathrm{d}\tilde{R}_F\right] \tag{B.36}$$

Now, the change in total welfare from a change in taxes satisfies

$$\mathrm{d}\mathcal{W} = \mathrm{d}V^N(0,X) - (1-N)(\mathrm{d}V^N(0,X) - \mathrm{d}\mathbb{E}_\theta V_E(0,\theta,X))$$

Substituting in (B.35) and (B.36), we obtain

$$\mathrm{d}\mathcal{W} = \frac{N e^N_{\tilde{w}}}{H'_z(z^*)\tilde{w}}\left[\mathrm{d}\tilde{w} + \mathcal{A}^N \mathrm{d}\tilde{R}_F\right] - \frac{(1-N)\mathrm{d}N}{H'_z(z^*)} = \frac{N e^N_{\tilde{w}}}{H'_z(z^*)\tilde{w}}\left[\mathrm{d}\tilde{w} + \mathcal{A}^N \mathrm{d}\tilde{R}_F - \frac{(1-N)\tilde{w}\mathrm{d}N}{N e^N_{\tilde{w}}}\right]$$

Using that $\tilde{w} = (1-\tau_N)w$, $\tilde{R}_F = (r_F - \delta)(1-\tau_K) - \tau_W$, $B^N_{\tau_K} = (r_F - \delta)\mathcal{A}^N$ and $B^N_{\tau_W} = \mathcal{A}^N$, it follows that the change in welfare from a change in $\tau_j \in \{\tau_K; \tau_W\}$, holding pre-tax prices constant, satisfies

$$\mathrm{d}\mathcal{W} = \frac{e^N_{\tilde{w}}}{H'_z(z^*)\tilde{w}}\left[-w\mathrm{d}\tau_N N - B^N_{\tau_j} N \mathrm{d}\tau_j - \frac{(1-N)w(1-\tau_N)\mathrm{d}N}{e^N_{\tilde{w}}}\right].$$

Now, under the assumptions of the Proposition, the government budget constraint is identical to the baseline model. Furthermore, since elasticities and aggregate variables are finite and non-zero, it follows that the budget constraint can be differentiated. Thus, it follows that, for a change in $\tau_j \in \{\tau_K; \tau_W\}$, holding pre-tax prices constant

$$0 = B_{\tau_N}\mathrm{d}\tau_N + B_{\tau_j}\mathrm{d}\tau_j + \sum_{,} \tau_m \frac{\partial B_{\tau_m}}{\partial \tau_j}\mathrm{d}\tau_j.$$

Combining this with the expression for the change in welfare above, we obtain

$$\mathrm{d}\mathcal{W} = \frac{e^N_{\tilde{w}}}{H'_z(z^*)\tilde{w}}\left[B_{\tau_j} + \left(\sum_{m\in\{K;W;N\}} \tau_m \frac{\partial B_{\tau_m}}{\partial \tau_j}\right) - B^N_{\tau_j} N - \frac{(1-N)w(1-\tau_N)}{e^N_{\tilde{w}}}\frac{\partial N}{\partial \tau_j}\right]\mathrm{d}\tau_j.$$

Then, as in the derivation of Proposition 1, the first order condition for the optimal choice of τ_j is that the resulting change in welfare from a small tax change is zero, so that

$$B_{\tau_j} + \left(\sum_{m \in \{K;W;N\}} \tau_m \frac{\partial B_{\tau_m}}{\partial \tau_j} \right) - B^N_{\tau_j} N - \frac{(1-N)w(1-\tau_N)}{e^N_{\tilde{w}}} \frac{\partial N}{\partial \tau_j} = 0,$$

which is the same first order condition as in the derivation of Proposition 1. Rearranging this to solve for optimal taxes and writing in matrix form, we arrive at the optimal tax formula in Proposition 1. □

B.8 Values of Terms in the Optimal Tax Formula

The remaining terms used in the optimal tax formula in Proposition 1 at the calibrated initial steady state are as follows

$$B = \begin{pmatrix} 0.12 & 0 \\ 0 & 3 \end{pmatrix}, \quad g_1 = \begin{pmatrix} -0.075 & -4.197 \\ -0.004 & -0.110 \end{pmatrix}.$$

C Data

To calibrate the entrepreneur's stake in the business, we use data from two sources: the Survey of Consumer Finances (SCF) and the (National) Survey of Small Business Finances (SSBF). Both surveys contain information regarding business ownership, with the difference that the first is a household survey, while the second is a survey of small businesses. We can identify in each of them groups of respondents that are in line with our notion of entrepreneurship. We use both sources as validation for our results.

The Survey of Consumer Finances is a a triennial cross-sectional survey of U.S. families which provides information on individual household portfolio composition, including investment in private firms. While the SCF was initially administered in 1983, it was not until 1989 that questions about business ownership were introduced. Therefore, we use all survey waves from 1989 until 2013. We restrict the sample to households who report owning a business in which they have an active management interest, and are between 25 and 65 years old. This represents, on average, 14.3% of the sample. If a household is an active participant in multiple businesses, we examine the average share across businesses.[32]

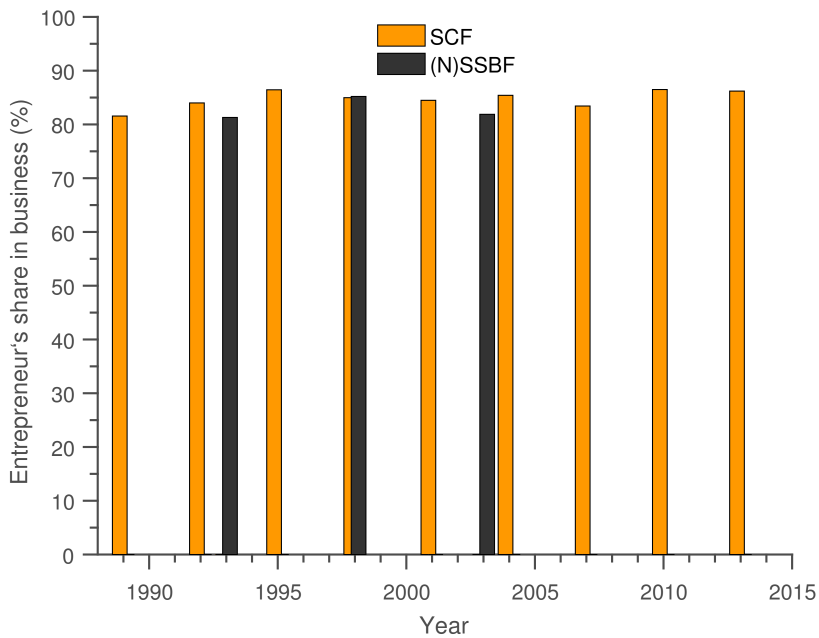
The (National) Survey of Small Business Finances collects information on private, non-financial, non-agricultural businesses in the U.S., with fewer than 500 employees. Only the surveys collected in 1993, 1998 and 2003 have ownership share information. The surveys detail the demographic and financial characteristics of the firms and their principal share-

[32] We obtain similar results if we focus on the business in which the household has the largest investment.

holder. Approximately 90% of these firms are managed by the principal shareholder. We apply the same sample restrictions as in the SCF.

Figure 4 displays the evolution of the ownership share over time. Both surveys indicate that ownership is highly concentrated, entrepreneurs holding, on average, 84% of their firm's equity. In particular, the average share is 85% in SCF and 83% in (N)SSBF. Ownership rates are very stable not only across surveys, but also across the time horizon we consider, so for our calibration exercise we work with their average over time and surveys, 84%.

Figure 4: Ownership Share in the U.S.



Notes: The orange bars show the average share that entrepreneurs in SCF own in their business. The black bars show the average share of small businesses in the (N)SSBF that is owned by the principal shareholder.