Optimal Taxation of Risky Entrepreneurial Capital

Corina Boar
New York University

Matthew Knowles
University of Cologne[*,†]

April 9, 2022

Abstract

We study optimal taxation in a model with endogenous financial frictions, risky investment and occupational choice, where the distribution of wealth across entrepreneurs 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 a redistributive motive for taxation. Optimal tax rates can be written as a closed-form function of the size of the tax bases and their elasticities with respect to tax rates. We find that it is optimal to tax capital income because financial frictions reduce the elasticity of capital income with respect to taxes and because capital income taxes prevent excessive entry into entrepreneurship. Optimal wealth taxes are positive but close to zero, since they strongly discourage capital accumulation.

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 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 and the National Tax Association's 113th Annual Conference on Taxation. 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: University of Cologne, email: `knowles@wiso.uni-koeln.de`

1 Introduction

This paper studies the optimal taxation of capital income and wealth in a model where taxation affects how efficiently capital is allocated in the economy. 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] Additionally, it is often argued that taxation may affect incentives for entrepreneurs to take risks, implying that taxation may affect the allocation of capital between more and less risky uses.[2]

We analyze optimal linear taxation in a model which incorporates these issues. In the model, there are overlapping generations of two types of households: workers and entrepreneurs. Newborn households decide whether to become workers or entrepreneurs, and retain the same job for their entire life. Entrepreneurs, who differ in ability, choose how much capital to allocate to a risky technology and to a risk-free technology. Furthermore, they lend to one another through frictional financial markets, where this 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 risks. This discourages them from allocating capital to the risky technology, which consequently has a higher expected return in equilibrium.

A utilitarian planner sets linear tax rates on capital income, wealth and labor income in order to maximize the lifetime utility of a newborn household in the long-run. Since entrepreneurial risks lead to significant inequality, the utilitarian planner has an incentive to redistribute between agents, but this entails efficiency costs. Namely, taxes on capital income and wealth affect how efficiently capital is allocated in the economy 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 choose to become entrepreneurs, which further affects the allocation of capital.

Our model is highly tractable and allows us to study many of these effects analytically, without imposing specific functional forms on the production technology or entrepreneurial ability distribution. Optimal tax rates in the model 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).

[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.

We characterize analytically the partial equilibrium elasticities that enter into the optimal tax formula. This allows us to study the forces that influence them, and, consequently, optimal tax rates. These partial equilibrium elasticities hold prices constant, but incorporate and greatly depend upon the long run endogenous response of the wealth distribution to tax rates. Capital income taxes are not equivalent to wealth taxes in our setting, unlike in traditional models, because the risky and risk-free technologies differ in their average return to capital. Therefore, capital income taxes fall relatively more on the high-return risky technology, while wealth taxes affect all capital equally. We find that capital income taxes have a relatively stronger tendency to inefficiently shift capital away from the high-return risky technology. This is because, first, capital income taxes reduce the post-tax excess return to the risky technology relative to the risk free technology, and so encourage entrepreneurs to shift capital towards the risk-free technology. Second, they fall most heavily on high ability entrepreneurs who earn a high rate of return to their wealth overall. As such, they tend to reduce the share of wealth of high ability entrepreneurs, which also shifts the allocation of productive capital away from these entrepreneurs due to financial frictions, as in Guvenen et al. (2019). Third, capital income taxes discourage entrepreneurial entry more than wealth taxes do, since workers only earn the lower risk-free return to their savings and so are less affected by capital income taxes. By decreasing the number of entrepreneurs, capital income taxes also shift capital away from the risky technology.

Our optimal tax formula reveals two main motivations for taxing capital income and wealth. First, a redistributive motive: much of capital income and wealth is earned by rich entrepreneurs with a low marginal utility of consumption, whereas labor income is earned by poorer agents. Second, low or negative capital income or wealth taxes increase entry into entrepreneurship, which reduces tax revenue if most taxes fall on workers. Thus, government revenue is increased by reducing entry into entrepreneurship, and capital income and wealth taxes are a valuable tool to achieve this. Importantly, in our model the long run elasticities of all tax bases with respect to taxes are generically finite. Thus, the model is not subject to the zero-optimal-tax results of Chamley (1986) and Judd (1985) which were derived in models with infinite long run elasticities (see Straub and Werning, 2020).

We calibrate the model to match values of tax bases, rates of return and features of financial contracts in the data. The calibrated model generates top wealth inequality similar to the data, as well as a large (7%) gap between risky and risk-free rates of return. In our baseline calibration, we find that it is optimal to tax capital income and wealth at low positive rates, with most taxes falling on labor income. The optimal capital income tax is 3.7%, the optimal wealth tax is 0.2% and the optimal labor income tax is 28%. Nevertheless, we find that the welfare gains from shifting from the current US tax policy to these optimal tax rates are small, a mere 0.2% of consumption equivalent variation in the steady state.

We analyze how optimal taxes and the elasticities underlying them depend on financial frictions by studying versions of the model with looser financial frictions as well as with exogenous financial frictions. Looser financial frictions tend to reduce optimal capital income taxes, while raising optimal wealth taxes. This is because looser financial frictions increase the tendency to which capital income represents a return to saving rather than economic profits, making it more elastic in response to taxation. Optimal taxes on capital income are higher when the endogenous financial friction in the model is replaced by a simple exogenous one. This is because, under the exogenous financial friction, the fraction of capital in the risky technology is less elastic with respect to the capital income tax rate.

Related literature. This paper studies optimal taxation in an environment in which output depends on the allocation of capital across heterogeneous entrepreneurs, which is affected by taxation. In that sense, our paper builds upon the related work of Evans (2015), Shourideh (2014), Itskhoki and Moll (2019), Guvenen et al. (2019), Boar and Midrigan (2020), Basseto and Cui (2020). We differ from these papers, first, by characterizing optimal taxes on capital income and wealth as closed form functions of "sufficient statistics", which not only enables us to carefully inspect the equity-efficiency tradeoff theoretically but also to provide a bridge between theory and empirics; second, by micro-founding the financial friction (thus allowing for changes in taxes to lead to changes in the tightness of financial frictions), and/or, third, allowing for individuals to endogenously choose whether to be an entrepreneur or worker. Within this literature, our paper is closest to Guvenen et al. (2019), who also focus on the different effects of capital income and wealth taxation on the allocation of capital through similar channels. Unlike us, they focus on numerical results and assume an exogenous financial friction. Differently from Guvenen et al. (2019), we find the optimal wealth tax is close to zero, whereas their result shows much higher optimal wealth taxes. This is because we allow for endogenous entry into entrepreneurship, which creates additional incentives to tax capital income rather than wealth.

Our paper is also related to the work that studies the effects of changing taxes numerically in models with entrepreneurs with heterogeneous productivity levels. Examples are Cagetti and De Nardi (2009), Kitao (2008), 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 and we focus on analytical rather than numerical results, with the aim of making the intuition behind the key mechanisms as transparent as possible and exploring the effects of a wider range of tax policy changes. Second, our financial friction arises endogenously as a consequence of asymmetric information between entrepreneurs and financial intermediaries and, as such, our results highlight that the degree to which financial markets are frictional may itself be affected by changes in taxes and that this is of importance when considering optimal taxation.

Our paper also relates to Panousi and Reis (2014), Panousi and Reis (2019) and Phelan (2019), who study optimal taxation in the presence of idiosyncratic investment risk. In these papers, unlike our setting, entrepreneurs do not differ in their expected productivity levels and there is only one production technology, so the allocation of capital does not itself affect aggregate output. Furthermore, the endogenous effects of tax changes on financial frictions in our setting substantially mitigates the role that capital income taxation can play to insure against investment risk in Panousi and Reis (2014) and Panousi and Reis (2019).

Lastly, our paper contributes to the wider literature on optimal capital taxation, which generally 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 (2019), Chen et al. (2019), among others.[3] Related to our paper, Abo-Zaid (2014), Biljanovska (2019) and Biljanovska and Vardoulakis (2019) have explored 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 derives properties of the model equilibrium and steady state and shows how the steady state is affected by tax rates. Section 4 derives formulae for the optimal tax rates and shows the values of optimal taxes in the numerical calibration. Section 5 concludes.

2 Model

In this section we describe our model economy and define an equilibrium. As we discuss in Section 4.3, our main result for optimal tax formulae does not depend on many of the details of the model, but we find it instructive to provide these details here for context.

Environment We consider a discrete time, infinite-horizon economy populated by a unit mass of households and a continuum of 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. Entrepreneurs manage firms and workers supply labor. Entrepreneurs use capital to produce intermediate goods. In particular, each entrepreneur is the owner of two different investment projects: a risky project which produces 'risky' intermediate goods denoted by y_E, and a risk-free project, which produces 'risk-free' intermediate goods denoted by y_F.[4] Entrepreneurs use labor in combination with

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

[4] The device of having two separate types of intermediate goods is a simple way to incorporate misallocation of capital into the model and to allow entrepreneurs to choose between allocating capital in a risky way or risk-free way. Alternatively, the risk free project can be interpreted as capital rented to a risk-free corporate sector which produces a different intermediate good to entrepreneurs.

intermediate goods, which they can trade among themselves, to produce a final good. The government levies taxes on households and funds exogenous government spending $\overline{G}$.

Timing Each period t is divided into three sub-periods: morning, afternoon and evening. In the morning, entrepreneurs buy and sell capital amongst themselves and each entrepreneur freely divides her capital between her risky and her risk-free investment projects. In the afternoon, each entrepreneur draws an idiosyncratic shock, which affects the quantity of capital in her risky project, and her two projects produce intermediate goods. Entrepreneurs sell the intermediate goods they produce amongst themselves. In the evening, entrepreneurs use intermediate goods and labor to produce the final good, which is sold to households. Households divide their resources between consumption and saving for the next period. At the end of the period, a fraction $\gamma \in (0,1)$ of households die and new households are born. Newborn households choose an occupation. Capital depreciates at rate $\delta \in (0,1)$.

Technology of Entrepreneurs At the beginning of each period t, each entrepreneur i is endowed with wealth $a_{i,t}$. In the morning, newborn entrepreneurs draw an ability level $\theta_{i,t} \in [0,1]$ from a continuous distribution with cdf $H_\theta(\cdot)$ and pdf $h_\theta(\cdot)$. At the start of each period, each continuing entrepreneur has the same θ as in the previous period with probability $1-\lambda_\theta$ and draws a new θ from the distribution $H_\theta(\cdot)$ with probability λ_θ.

After allocating capital between her risky and risk-free projects in the morning, the entrepreneur draws a idiosyncratic shock $\xi_{i,t} \in \mathbb{R}$ in the afternoon from the continuous distribution $H_\xi(\cdot)$ with pdf $h_\xi(\cdot)$. We assume that H_ξ has mean zero, standard deviation 1 and full support on $\mathbb{R}$.[5] Therefore, an entrepreneur with type $\theta_{i,t}$ who allocates $k_{E,i,t}$ to her risky project in the morning and draws shock $\xi_{i,t}$ has $\tilde{k}_{E,i,t} = q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t})$ units of capital in her risky project in the afternoon, where $q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t})$ is increasing in $k_{E,i,t}$ and $\xi_{i,t}$. Each unit of capital $\tilde{k}_{E,i,t}$ in the risky project produces one unit of risky intermediate goods. Thus, the entrepreneur's output of these goods $y_{E,i,t}$ satisfies $E[y_{E,i,t}] = E[\tilde{k}_{E,i,t}] = k_{E,i,t}$. The risk-free project produces an output of $y_{F,i,t} = k_{F,i,t}$ risk-free intermediate goods.

We assume the following functional form for $q(\theta_{i,t}, \xi_{i,t}, k_{E,i,t})$

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

where $\underline{\epsilon} \in (0,1)$, $\varphi > 0$ and $\underline{k}_E > 0$. This functional form implies that an entrepreneur can put $k_{E,i,t} \leq \underline{k}_E$ into her risky project without facing any risk at all. If $k_{E,i,t} > \underline{k}_E$ then the entrepreneur's risky project genuinely becomes risky. In this case the mean of $\tilde{k}_{E,i,t}$ is equal

[5] These restrictions on the first two moments of $H_\xi(\cdot)$ and the upper bound on θ are purely normalizations.

to $k_{E,i,t}$, and the variance of $\tilde{k}_{E,i,t}$ is decreasing in $\theta_{i,t}$.[6] The lowest possible realization of $\tilde{k}_{E,i,t}$ is $(1-\underline{\epsilon})k_{E,i,t} + \underline{\epsilon k}_E$. Thus, the advantage of having higher $\theta_{i,t}$ is that it reduces the risk that entrepreneurs face for a given amount of capital in the risky project.

Our assumption that entrepreneurial production is subject to idiosyncratic risks follows a large literature.[7] Our specific assumption on the functional form of $q(\cdot)$ is not critical for our derivation of the optimal tax formula, as we discuss in Section 4.3. However, it has a number of convenient properties. First, it implies that optimal contracts for obtaining external funds are identical to equity and debt contracts, which are the most common financial contracts in the data. This helps with calibrating the agency frictions against the data. Second, it allows us to study the effects of taxes on the allocation of capital and on entrepreneurial entry in a relatively tractable way, while allowing for heterogeneous abilities across entrepreneurs.

Hiding Capital Entrepreneurs are able to hide capital $k_{H,i,t}$ in their risky project after observing their shock $\xi_{i,t}$ and convert it directly 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 of entrepreneurs to hide capital affects allocations and optimal taxes by creating frictions in financial markets.

Technology of Final Good Production Entrepreneurs trade risky intermediate goods at price $r_{E,t}$ per unit and risk-free intermediate goods at price $r_{F,t}$ per unit. These prices are the expected return to risky projects and the market rate of return to capital in risk-free projects, respectively. Each entrepreneur i hires $n_{i,t}$ workers at wage rate w_t and uses $y^d_{E,i,t}$ and $y^d_{F,i,t}$ units of risky and risk-free intermediate goods to produce $y_{i,t}$ final goods according to the production function $y_{i,t} = f\left(y^d_{E,i,t}, y^d_{F,i,t}, n_{i,t}\right)$, where f is concave and strictly increasing in all arguments, exhibits constant returns to scale and satisfies the Inada conditions.

Preferences Each worker has a constant labor endowment equal to 1, supplied inelastically. The consumption of an entrepreneur i is denoted by $c_{i,t}$ and that of a worker i is denoted by $c^N_{i,t}$. 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 agent i's period utility in period $t+j$. For a worker, $u_{i,t+j}$ is equal to $\log(c^N_{i,t+j})$. For an entrepreneur, $u_{i,t+j}$ is equal to $\log(c_{i,t+j}) + z_i$, where $z_i \in \mathbb{R}$ is an individual-specific utility of being an entrepreneur, which can be viewed as representing an individual's taste for the non-pecuniary benefits (e.g. the pleasure of 'being one's own boss') that Hurst and Pugsley (2011) find are an important motivation in individuals choosing to become entrepreneurs.[9]

[6]The variance of $\tilde{k}_{E,i,t} - \underline{k}_E$ is inversely proportional to $\theta_{i,t}$. This is without loss of generality.

[7]For instance, Bernanke et al. (1999) and Panousi (2012).

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

[9]Allowing some individuals to value the non-pecuniary benefits of entrepreneurship helps the model to match the number of entrepreneurs in the data.

At birth, individuals draw their time-invariant value of z_i from the continuous probability distribution $H_z(\cdot)$ with pdf $h_z(\cdot)$. We impose the restriction that $\varphi^2 > \lambda_\theta + \rho + 2\gamma$, which holds easily in our calibration and guarantees that, in a steady state, the allocation of capital to risky projects is proportional to entrepreneurial ability, thus simplifying aggregation.

Occupational Choice Newborn households choose their occupation to maximize expected lifetime utility. They are identical except for their value of z_i, so there exists a cutoff $z_t^\star$ such that individuals born at time t with $z_i \geq z_t^\star$ will choose to be entrepreneurs and individuals with $z_i < z_t^\star$ will choose to be workers. The cutoff $z_t^\star$ satisfies

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

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

Government The government levies three different taxes: 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 $\overline{G}$, while balancing its budget every period. Taxes are paid in the evening and government spending also takes place in the evening. The government's budget constraint each period is

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

where N_t is the total measure of workers, K_t is the aggregate capital stock at the start of the period and $\Pi_t - \delta K_t$ is the total reported profits of entrepreneurs net of capital depreciation.[10]

Financial Markets Entrepreneurs may fund capital purchases by writing one-period state-contingent financial contracts with risk neutral and perfectly competitive banks, that live for only one period. An entrepreneur who borrows some quantity $b_{i,t} > 0$ in the morning returns quantity $\hat{b}_{i,t}$ at the end of the period. It is convenient to write the entrepreneur's choices of $b_{i,t}$ and $\hat{b}_{i,t}$ as policy functions of the relevant state variables. In general, an entrepreneur's choice of $b_{i,t}$ will depend on her ability $\theta_{i,t}$, her start of period wealth $a_{i,t}$ and the aggregate state of the economy, which we label X_t. Therefore, abusing notation slightly, we write $b_{i,t} \equiv b(a, \theta, X)$ and $\hat{b}_{i,t} \equiv \hat{b}\,(a, \theta, \xi, X)$.

A bank will only lend to entrepreneurs in the morning if the expected return on the loan in the evening is equal to the market risk-free rate. This implies the following constraint

$$\int_\xi \hat{b}\,(a, \theta, \xi, X)\, \mathrm{d}H_\xi\,(\xi) \geq R_{F,t} b\,(a, \theta, X)\,,$$

[10] As will be seen below, entrepreneurs have private information about the return they earn on their capital. Since the government is not able to observe this private information , the government's tax revenue depends on entrepreneurs' *reported* (rather than realized) profits.

where $R_{F,t}$ denotes the gross market risk-free rate of interest within the period. In equilibrium, this inequality will be satisfied with equality and banks make zero profits. Workers can also borrow or lend to banks within the period at the risk free rate $R_{F,t}$.

Annuities At the end of the period households may trade among themselves financial annuities, which insure against the risk of death. A household may exchange a unit of the final good at the end of the period for the promise of receiving $\frac{1}{1-\gamma}$ units of the final good at the start of the next period as long as the household is still alive. Entrepreneurs place all their capital in a common fund at the end of the period, exchanging it for annuities.[11]

Budget Constraints The worker's budget constraint is

$$c^N_{i,t} + (1-\gamma)a^N_{i,t+1} \quad = \quad w_t(1-\tau_{N,t}) + R_{F,t}a^N_{i,t},$$

where $a^N_{i,t}$ denote the worker's start of period assets.[12]

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 in the risky project to hide, $k_{H,i,t}$. In the evening, she chooses consumption $c_{i,t}$ and annuities $(1-\gamma)a_{i,t+1}$. Capital hidden in the afternoon is transformed into $c_{H,i,t}$ units of consumption. Finally, in the evening the entrepreneur repays the bank $\hat{b}_{i,t}$ and pays her taxes to the government. Consequently, in the evening the entrepreneur's budget constraint is

$$c_{i,t} - c_{H,i,t} + (1-\gamma)a_{i,t+1} + \hat{b}_{i,t} \quad = \quad \pi_{i,t} - T_{i,t} + (1-\delta)\left((\tilde{k}_{E,i,t} - k_{H,i,t}) + k_{F,i,t}\right),$$

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

$$\pi_{i,t} \quad = \quad \underbrace{(r_{E,t}y_{E,i,t} + r_{F,t}y_{F,i,t})}_{\text{profit from intermediate goods}} + \underbrace{\left(y_{i,t} - w_t n_{i,t} - r_{E,t}y^d_{E,i,t} - r_{F,t}y^d_{F,i,t}\right)}_{\text{profit from final good}},$$

[11]If entrepreneurs were allowed to hold capital rather than placing in the common fund, they would prefer the common fund, since this insures against the risk associated with the stochastic death.

[12]Note that workers do not pay capital income or wealth taxes. For analytical convenience, we assume such taxes are levied on physical assets and entrepreneurial profits only, rather than on financial assets. This is without loss of generality. If linear tax rates were also imposed on net financial positions and net financial wealth, these taxes would simply lead to an adjustment of pre-tax interest rates, and leave the post-tax rate of interest and allocations unaffected, as in the textbook discussion in Varian (2014), p. 307.

$T_{i,t}$ is the entrepreneur's period tax payments given by[13]

$$T_{i,t} = \tau_{K,t}\pi_{i,t} - \tau_{K,t}\delta k_{i,t} + \tau_{W,t}k_{i,t},$$

and where $c_{H,i,t}$, $y_{E,i,t}$, $y_{F,i,t}$ satisfy

$$c_{H,i,t} = \phi k_{H,i,t} \qquad y_{E,i,t} = \tilde{k}_{E,i,t} \qquad y_{F,i,t} = k_{F,i,t}.$$

Agency Friction An entrepreneur's realization of $\xi_{i,t}$, the capital in the risky project that she hides and the consumption she obtains from converting hidden capital are all private information.[14] Without loss of generality, we restrict attention to incentive compatible contracts with banks where the entrepreneur honestly reports her $\xi_{i,t}$ and pays the promised amount $\hat{b}(a, \theta, \xi, X)$. This gives rise to the following incentive compatibility constraint

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

Under-reporting ξ by a small amount $\mathrm{d}\xi$ entails the entrepreneur hiding $\frac{\partial \tilde{k}_E}{\partial \xi}\mathrm{d}\xi$ units of capital, so she can produce $\frac{\partial \tilde{k}_E}{\partial \xi}\mathrm{d}\xi$ fewer units of intermediate goods from the risky technology and loses the after-tax return from selling them. At the same time she transforms 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.[15]

Worker's Optimization Problem The worker chooses consumption $c^N(a^N, X)$ and annuities $a^{N\prime}(a^N, X)$ that solve the following Bellman equation

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

subject to the worker's budget constraint.

Entrepreneur's Optimization Problem As discussed above, the expected present

[13]Note that entrepreneurs do not pay any tax on the capital gain $\tilde{k}_{E,i,t} - k_{E,i,t}$. This is without loss of generality for two reasons. First, capital gains are zero on average so the tax does not generate government revenue. Second, a moderate tax on this capital gain has no effect on equilibrium allocations because a consequence of our endogenous financial friction is that banks absorb all idiosyncratic risk beyond the minimum required for incentive compatibility, which is unaffected by taxes.

[14]In the extreme case $\phi = 0$ there would be no informational friction, since the entrepreneur has no incentive to hide capital.

[15]As we discuss below, this agency friction leads in equilibrium to restrictions on entrepreneurs' ability to obtain external finance (Bernanke et al., 1999). Since these financial frictions arise endogenously, the level of external finance entrepreneurs can obtain may be affected by taxation. The specific agency friction assumed here allows the effect of taxes on financial frictions to be studied in a relatively tractable way.

discounted lifetime utility of the entrepreneur i born at time t is

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

where $V(a,\theta,X_t)$ is the continuation value of an entrepreneur with assets a and type θ, ignoring the $\frac{z_i}{1-(1-\rho)(1-\gamma)}$ term. The entrepreneur chooses non-negative functions k_E, k_F, k_H, b, $\hat{b}$, c, c_H, y_E, y_F, y_E^d, y_F^d, n, y, as well as a possibly negative function a' to solve

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

subject to the mornign 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.

Aggregation and Market Clearing Aggregate consumption C_t and of $C_{H,t}$ satisfy

$$C_t=\int_{i\in\mathcal{N}_t}c_{i,t}^N\mathrm{d}i+\int_{i\notin\mathcal{N}_t}c_{i,t}\mathrm{d}i \qquad \text{and} \qquad C_{H,t}=\int_{i\notin\mathcal{N}_t}c_{H,i,t}\mathrm{d}i\leq C_t.$$

where $\mathcal{N}_t$ is the set of workers at time t: $\mathcal{N}_t:=\bigcup_{s\leq t}\{i: i \text{ is born in period } s \text{ and } z_i<z_s^\star\}$.

Aggregate capital devoted to each use and aggregate final good output are given by

$$K_{E,t}=\int_{i\notin\mathcal{N}_t}k_{E,i,t}\mathrm{d}i,\qquad K_{F,t}=\int_{i\notin\mathcal{N}_t}k_{F,i,t}\mathrm{d}i,\qquad K_{H,t}=\int_{i\notin\mathcal{N}_t}k_{H,i,t}\mathrm{d}i,\qquad Y_t=\int_{i\notin\mathcal{N}_t}y_{i,t}\mathrm{d}i$$

Total reported period profits of entrepreneurs are $\Pi_t=Y_t-w_tN_t$. In each period, the asset market must clear, so the total capital stock equals the wealth of entrepreneurs and workers

$$\int_{i\in\mathcal{N}_t}a_{i,t}^N\mathrm{d}i+\int_{i\notin\mathcal{N}_t}a_{i,t}\mathrm{d}i \quad=\quad K_t=K_{E,t}+K_{F,t}.$$

The market for intermediate goods of each type must clear each period

$$\int_{i\notin\mathcal{N}_t}y_{E,i,t}^d\mathrm{d}i=\int_{i\notin\mathcal{N}_t}y_{E,i,t}\mathrm{d}i \qquad\text{and}\qquad \int_{i\notin\mathcal{N}_t}y_{F,i,t}^d\mathrm{d}i=\int_{i\notin\mathcal{N}_t}y_{F,i,t}\mathrm{d}i.$$

The labor market must clear each period

$$\int_{i\notin\mathcal{N}_t}n_{i,t}\mathrm{d}i \quad=\quad \int_{i\in\mathcal{N}_t}\mathrm{d}i=N_t$$

The final goods market clearing condition then follows by Walras' law

$$\bar{G}+C_t+K_{t+1}\quad=\quad Y_t+(1-\delta)K_t-(1-\delta)K_{H,t}+C_{H,t}.$$

Equilibrium We are now in the position to define an equilibrium for our model economy.

Definition 1. *Given a sequence of tax rates* $\{\tau_{W,t}, \tau_{K,t}, \tau_{N,t}\}_{t=0}^{\infty}$*, an equilibrium* $\mathcal{E}$ *is a sequence of prices* $\{R_{F,t}, r_{E,t}, r_{F,t}, w_t\}_{t=0}^{\infty}$*, decision rules of entrepreneurs and workers, and a sequence of aggregate variables* $\{C_t, C_{H,t}, K_t, K_{E,t}, K_{F,t}, K_{H,t}, Y_t, N_t\}_{t=0}^{\infty}$ *such that:*

1. *The government's budget constraint is balanced every period.*
2. *Workers' decision rules solve the worker's optimization problem.*
3. *Entrepreneurs' decision rules are given by the solution to the entrepreneur's problem.*
4. $\{C_t, C_{H,t}, K_t, K_{E,t}, K_{F,t}, K_{H,t}, Y_t\}_{t=0}^{\infty}$ *are aggregates of households' decisions defined above.*
5. *Newborn agents choose the occupation that maximizes expected lifetime utility.*
6. *The asset, intermediate goods and labor markets clear.*

3 Properties of the Model Equilibrium

In this section, we characterize the equilibrium and derive results for how aggregate variables change in response to changes in taxes, which we then use in characterizing optimal taxes.

3.1 Worker's Optimal Decisions

In Appendix A.1, we show that the solution to the worker's problem is given by

$$c_t^N = [1-(1-\rho)(1-\gamma)]\, R_{F,t} P_t^N \qquad \text{and} \qquad P_{t+1}^N = (1-\rho)\, R_{F,t} P_t^N,$$

where $P_t^N \equiv a_t^N + \underbrace{\sum_{j=0}^{\infty}\left[\frac{w_{t+j}(1-\tau_{N,t+j})(1-\gamma)^j}{\Pi_{k=0}^{j} R_{F,t+k}}\right]}_{F_t^N}$ denotes the discounted value of lifetime income. The worker hence devotes a constant fraction of her discounted lifetime income to consumption expenditure and saves the remainder. The the associated value function is

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

3.2 Entrepreneur's Optimal Decisions

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 end-of-period resources between c and a'.

To that end, we first note that since final goods production has constant returns to scale, all the profits that the entrepreneur makes accrue from selling intermediate goods, so $\pi_{i,t}$ is

$$\pi_{i,t} = r_{E,t}(\tilde{k}_{E,i,t} - k_{H,i,t}) + r_{F,t}k_{F,t}.$$

Next, we define $P_{i,t}$ as present value of the lifetime resources obtained by an entrepreneur who takes no risk. 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 $a_{i,t} - \underline{k}_E$ to banks at the risk-free rate $R_{F,t}$. Therefore, $P_{i,t}$ is

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

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

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

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[16]

$$\tilde{V}(\omega, \theta, X) = \sup_{c, P'} \Big(\log(c) + (1-\rho)\,(1-\gamma)\mathbb{E}V(P', \theta', X') \Big), \tag{3}$$

$$\text{s.t.} \quad c + (1-\gamma)\,P' = \omega. \tag{4}$$

The entrepreneur's within-period problem is to choose non-negative functions $k_E(P, \theta, X)$, $k_F(P, \theta, X)$, $k_H(P, \theta, X)$, $\omega(P, \theta, \epsilon, X)$ and functions $b(P, \theta, X)$, $\hat{b}\,(P, \theta, \epsilon, X)$ to solve

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

[16] We can relabel $V\,(a, \theta, X)$ as $V\,(P, \theta, X)$ since $P = a + F$ and F depends only on the aggregate state X.

$$
\begin{aligned}
\text{s.t.}\quad b &= k_E + k_F - P + F && (5)\\
0 &= \int_{\xi} (\hat{\hat{b}} - R_F b)\mathrm{d}H_{\xi}(\xi) && (6)\\
\omega &= [\phi - (1-\tau_K) r_E - (1-\delta)]k_H - (\hat{b} - R_F b) + R_F P \\
&\quad + [(1-\tau_K)\, r_E + (1-\delta)](\tilde{k}_E - \underline{k}_E) + (\tau_K \delta - \tau_W - R_F)\,(k_E - \underline{k}_E) \\
&\quad + [-R_F + 1 + (1-\tau_K)\,(r_F - \delta) - \tau_W]k_F && (7)\\
k_H &\leq k_E && (8)\\
\frac{\partial \omega}{\partial \xi} &\geq \phi \frac{\partial \tilde{k}_E}{\partial \xi}. && (9)
\end{aligned}
$$

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

$$
\begin{aligned}
R_{F,t} &= 1 + (1-\tau_{K,t})\,(r_{F,t} - \delta) - \tau_{W,t}, && (10)\\
\phi &\leq (1-\tau_{K,t}) r_{E,t} + (1-\delta), \text{ with equality if } k_{H,i,t} > 0, && (11)\\
0 &< (r_{E,t} - r_{F,t})(1-\tau_{K,t}). && (12)
\end{aligned}
$$

Equation (10) states that the risk-free return to lending to a bank equals the return to putting capital in the risk-free technology. Equation (11) states that hiding capital cannot be more lucrative than selling risky intermediate goods. Equation (12) implies that the risky technology has a higher return than the risk-free technology, to compensate for risk. In equilibrium, each entrepreneur chooses $k_{E,i,t} \geq \underline{k}_{E,i,t}$, since borrowing and investing up to $\underline{k}_{E,i,t}$ is possible without risk, and equation (12) implies that this yields a positive return.

Our assumptions on the function $q(\cdot)$ 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 the following Lemma.

Lemma 1. *Let $\overline{V}(\theta, X) = V(1, \theta, X)$, then, for any P, θ and X,*

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

Proof. See Appendix A.2. □

The solution to the between period problem is then easily shown to be

$$
c = (1 - (1-\rho)\,(1-\gamma))\,\omega \qquad \text{and} \qquad P' = (1-\rho)\,\omega.
$$

To solve the within period problem, note that the risk averse entrepreneur chooses a contract that minimizes the variance of ω while satisfying the incentive compatibility constraint (9). Integrating with respect to ξ, it follows that there exists a function $\underline{\omega}(P,\theta,X)$ such that

$$\omega(P,\theta,\xi,X) \equiv \underline{\omega}(P,\theta,X) + \phi(q(\theta,\xi,k_E(P,\theta,X)) - \underline{k}_E), \tag{13}$$

which we write more compactly as $\omega \equiv \underline{\omega} + \phi(\tilde{k}_E - \underline{k}_E)$.

In the absence of agency frictions, the entrepreneur and the 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 (13). This uniquely pins down the value of $\hat{b}(P,\theta,\xi,X)$ 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 2. *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 $1 - \frac{\phi}{r_E(1-\tau_K)+(1-\delta)}$ of the remaining value of her risky project as equity, retaining the fraction $\frac{\phi}{r_E(1-\tau_K)+(1-\delta)}$ of the equity herself.*

Proof. See Appendix A.3 □

The reason the entrepreneur cannot sell all the equity in her project 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.

Combining the rewritten incentive compatibility constraint (13) with the definition of ω, integrating with respect to ξ and using equations (10)-(11) reveals that $\underline{\omega}(\cdot)$ 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:

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

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

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.

3.3 Continuous Time Limit

To describe the environment, the equilibrium conditions and to simplify the contracting problem between the entrepreneurs and banks it was natural to make the assumption of discrete time. To characterize the steady state and solve for the optimal taxes, it is more convenient to work in continuous time. We formally derive a continuous-time version of our discrete-time economy in Appendix B. Solving the optimization problems of the entrepreneur and worker yields the following optimal decision rules.

Proposition 1. *In equilibrium, the unique solution of the worker's problem is*

$$\begin{aligned} c^N &= (\rho+\gamma)P^N \\ dP^N &= \left\{\left[\tilde{R}_F+\gamma\right]P^N - c^N\right\}dt, \end{aligned}$$

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

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

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

$$\hat{k}_E(\theta) \equiv \frac{1}{\phi(1-\underline{\epsilon})} \times \min\left[\frac{(r_E-r_F)(1-\tau_K)\theta}{\phi(1-\underline{\epsilon})\varphi^2};1\right]. \tag{14}$$

Proof. See Appendix B.2 and Appendix B.3. □

Entrepreneurs do not hide capital in equilibrium because the return is always lower than selling risky intermediate goods.[17] All else equal, richer entrepreneurs invests more in risky projects. Furthermore,they invest more in risky projects and less in risk-free projects when: *(i)* the after-tax return to risky projects is relatively higher, *(ii)* the after-tax return to risk-free projects is relatively lower or *(iii)* the agency friction is less severe (i.e. lower ϕ).[18]

[17] As shown in Appendix B.3, the inequality (11) is strict.

[18] An implication is that capital income taxes reduce the willingness of entrepreneurs to invest in risky projects, which arises as a consequence of our endogenous financial frictions.

The aggregate value of capital in the risky technology satisfies

$$K_E = \underline{k}_E(1-N) + \mathbb{P}\int_0^1 \hat{k}_E(\theta)\mu(\theta)\mathrm{d}\theta \tag{15}$$

where $\mathbb{P}$ represents the aggregate of lifetime resources P and where $\mu(\theta)\mathrm{d}\theta$ denotes the share of lifetime resources that is held by entrepreneurs of type θ. Thus, the function $\mu(\cdot)$ represents the distribution of wealth across entrepreneurs. As such, the allocation of capital in the economy depends on the wealth distribution across entrepreneurs.

Equilibrium Occupational Choice Newborn households choose to be entrepreneurs iff they draw a level of z_i greater than $z^\star$. Therefore, $z^\star$ must satisfy

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

where the left hand side is the value of a newborn worker and the right hand side is the value of a newborn entrepreneur with $z_i = z^\star$. The total measure of workers evolves according to

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

3.4 Aggregate Steady State

We next formally characterize a steady state of the model.

3.4.1 Post-tax Prices

Equilibrium prices r_E, r_F and w are equal to the derivatives of $f(K_E, K-K_E, N)$ with respect to its three arguments. The planner can use the three tax instruments, τ_K, τ_W and τ_N to target the values of three post-tax prices, subject to the government budget constraint. As a consequence, the analysis of optimal taxation is made conceptually much easier by solving for agents' optimal choices in terms of post-tax prices and assuming that the government optimally chooses these prices by adjusting taxes accordingly. The three post-tax prices that determine agents' choices in this economy are $\tilde{R}_F$, defined above, the post-tax wage $\tilde{w} \equiv w(1-\tau_N)$ and post-tax excess return to the risky project $\tilde{r}_X \equiv (r_E - r_F)(1-\tau_K)$.

3.4.2 Steady State Characterization

We now formally define and characterize the steady state of the model.

Definition 2. *A* ***steady state*** $\mathscr{S}$ *of the economy is a set of tax rates* $\{\tau_W^*, \tau_K^*, \tau_N^*\}$*, prices* $\{r_E^*, r_F^*, w^*\}$*, aggregate variables* $\{K^*, K_E^*, C^*, N^*\}$ *and an equilibrium* $\mathcal{E}$ *in which all tax rates, prices and aggregates are equal to the steady state values in every period.*

The set of conditions that characterize a steady state are summarized in Proposition 2.

Proposition 2. *There exists a steady state $\mathscr{S}$ 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*

$$C^* = \frac{\rho+\gamma}{\tilde{R}_F^*+\gamma}\left(Y^* - \delta K^* - \overline{G} - \tilde{r}_X^*(K_E^* - (1-N^*)\underline{k}_E) + \gamma K^*\right) \tag{16}$$

$$C^* = Y^* - \delta K^* - \overline{G} = N^*\tilde{w}^* + \tilde{r}_X^* K_E^* + \tilde{R}_F^* K^* \tag{17}$$

$$N^* = H_z\left(\tilde{w}^* - \log \tilde{r}_x^* \underline{k}_E - \frac{\mathbb{E}[\theta]}{\rho+\gamma}\left(\tilde{r}_X^*\hat{k}(1) - \frac{(\phi(1-\underline{\epsilon})\hat{k}(1)\varphi)^2}{2}\right)\right) \tag{18}$$

$$K_E^* = \underline{k}_E(1-N^*)\left(1 + \frac{\frac{\gamma}{\gamma+\tilde{R}_F^*}\tilde{r}_X^*\int \hat{k}_E(\theta)\mu(\theta)\,d\theta}{\rho+\gamma-\tilde{R}_F^* - \tilde{r}_X^*\int \hat{k}_E(\theta)\mu(\theta)\,d\theta}\right) \tag{19}$$

$$\mu(\theta) = \frac{h_\theta(\theta)}{1 - \frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}}\left(\int_0^1 \frac{h_\theta(\theta)}{1 - \frac{\tilde{r}_X^*\hat{k}_E(\theta)}{\lambda_\theta+\rho+\gamma-\tilde{R}_F^*}}\,d\theta\right)^{-1} \tag{20}$$

$$\hat{k}_E(\theta) = \frac{\tilde{r}_X^*\theta}{(\phi(1-\underline{\epsilon})\varphi)^2} = \hat{k}_E(1)\theta \tag{21}$$

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

Proof. See Appendix B.4. □

These conditions are intuitive. The first is the consumption function. The marginal propensity to consume (MPC) is $\frac{\rho+\gamma}{\tilde{R}_F^*+\gamma}$ and agents consume from their net resources after tax and depreciation, $Y^* - \delta K^* - \overline{G}$ and the after-tax resources from annuities, γK^*. Entrepreneurs do not consume out of the risky part of their capital income, $K_E^* - (1-N^*)\underline{k}_E$, as a higher risky capital income means a higher return to saving. Equation (17) is the aggregate resource constraint. Equation (18) is the condition for occupational choice, where the argument of $H_z(\cdot)$ is the value of the cutoff $z^\star$. Equation (19) comes from equation (15) for K_E. In equations (19) and (20) we have solved for the stationary wealth distribution induced by entrepreneurs' policy choices which, as equation (21) shows, are linear in θ. For given parameter values, functional form assumptions for H_θ and H_z, and post-tax prices, it is straightforward to use these conditions to solve for the steady state numerically. We do this in Section 4.4 below.

3.5 The Effect of Taxes on Equilibrium Allocations

In Section 4.2 we show that optimal steady state tax rates can be written in terms of the partial equilibrium elasticities of aggregates with respect to taxes, in the spirit of Chetty

(2008), Piketty and Saez (2013), Piketty et al. (2014) and Saez and Stantcheva (2018). Here we formally define these elasticities and use them to gain intuition about the effect of the various tax instruments on key aggregates. To study the partial equilibrium effect of taxes on Y, K and K_E, we consider the long-run effect of a change in tax rates holding constant pre-tax prices, but allowing for variation over time in the distribution of wealth. We formally define long-run elasticities as follows.

Definition 3. *Let X^* be some aggregate variable, in the steady state. The partial equilibrium elasticities of X with respect to the tax rates τ_K and τ_W are defined as*[19]

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

By considering small perturbations in post-tax prices around the steady-state conditions in Proposition 2, the elasticities can be characterized in closed form, as functions of parameters and aggregate variables. We relegate the derivations to Appendix B.5 and provide the characterization of the elasticities of Y, K_E and K with respect to taxes below.

Effects of Tax Changes on Y

Proposition 3. *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*

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

This shows that a change in, for instance, τ_K affects aggregate output via its effect on aggregate capital accumulation ($e^K_{\tau_K}$), the fraction of agents who become workers ($e^N_{\tau_K}$) and the fraction of capital allocated to the risky technology ($e^{K_E}_{\tau_K}$ - $e^K_{\tau_K}$). The first two effects are the same as in a standard neoclassical growth model. The third effect arises because of financial market frictions, which imply that $r_E > r_F$ in the steady state, so an increase in the fraction of capital allocated to the risky technology increases output. 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.

Effects of Tax Changes on K_E

Proposition 4. *The partial equilibrium elasticity of steady state capital in the risky technology stock K_E with respect to the tax rates τ_j, $j \in \{K, W\}$, adjusting τ_N to balance the*

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

government's budget is

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

where

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

and

$$e_{\tau_K}^{\hat{k}_E} = (1-\tau_K)\frac{\partial}{\partial \tau_K}\log\left(\int_\theta \mu(\theta)\hat{k}_E(\theta))\,d\theta\right),$$

with an analogous definition of $e_{\tau_W}^{\hat{k}_E}$.

A change in τ_K or τ_W has three effects on K_E, captured by $e_{\tau_j}^{\hat{k}_E}$, $e_{\tau_j}^{\mathbb{P}}$ and $e_{\tau_j}^{N}$. The $e_{\tau_j}^{\hat{k}_E}$ term reflects that a tax change affects the average willingness of entrepreneurs to put capital into the risky technology by changing their choice of $\hat{k}_E$ and the distribution of wealth across entrepreneurs, which affects the weights $\mu(\theta)$. The term $e_{\tau_j}^{\mathbb{P}}$ 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 their income and encouraging consumption. Ultimately, this reduces their capital to put into the risky technology, so $e_{\tau_j}^{K_E}$ depends on $e_{\tau_j}^{\mathbb{P}}$. 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. The term $e_{\tau_j}^{N}$ 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 $e_{\tau_K}^{\hat{k}_E}$ and $e_{\tau_W}^{\hat{k}_E}$ are particularly relevant for how taxes affect the allocation of capital, since they determine how entrepreneurs allocate their capital to the risky technology. As the following Lemma shows, these elasticities are approximately -1 and 0 respectively, provided λ_θ is sufficiently large – a condition which is easily satisfied in our calibration below.

Lemma 3. *In the limit as λ_θ approaches infinity, the elasticities $e^{\hat{k}_E}_{\tau_K}$ and $e^{\hat{k}_E}_{\tau_W}$ satisfy*

$$e^{\hat{k}_E}_{\tau_K} \to -1,$$
$$e^{\hat{k}_E}_{\tau_W} \to 0.$$

Proof. See Appendix B.5. □

This indicates that capital income taxes have important effects on the allocation of capital, which wealth taxes do not. This is for two reasons. First, higher taxes on capital income reduce the post-tax excess return to the risky technology $\tilde{r}_X$, which directly reduces $\hat{k}_E(\theta)$ and shifts capital away from the risky technology and towards the risk-free technology. Second, capital income taxes fall relatively more strongly on entrepreneurs, particular those with high ability levels, who earn a higher return to capital.

Effects of Tax Changes on K

Proposition 5. *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*

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

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

The numerator in $e^K_{\tau_j}$ shows the degree to which a tax increase raises aggregate net saving, given K, and the denominator shows how aggregate saving decreases as K rises. The first term in the numerator captures the effect of a tax change on K_E. An increase in K_E increases aggregate saving via increasing entrepreneurs' income in proportion to $(r_E - r_F)(1 - \tau_K MPC)$. The second term $e^{S_{SUB}}_{\tau_j}$ captures the substitution effect induced by taxes directly affecting the post-tax return to saving. Finally, the third term captures the effect through N, as a higher fraction of workers, given K_E, increases saving by workers in proportion to $w(1 - MPC)$ and reduces consumption by entrepreneurs in proportion to $\tilde{r}_X \underline{k}_E$.

Effects of Tax Changes on N

We now derive the effect of a change in taxes on N. We present the algebraic derivation here, since we use similar arguments in deriving the optimal taxes below. 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 (18) with respect to τ_W, we obtain that

$$\frac{\partial N}{\partial \tau_W} = \frac{H_z'(z^\star)}{\tilde{w}}\frac{\partial \tilde{w}}{\partial \tau_W} = -\frac{H_z'(z^*)}{N}\frac{N}{1-\tau_N}\frac{\partial \tau_N}{\partial \tau_W} = \frac{-e_{\tilde{w}}^N N}{1-\tau_N}\frac{\partial \tau_N}{\partial \tau_W}, \tag{22}$$

where we used that $z^\star$ is the argument of the cdf $H_z(\cdot)$ in equation (18) 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

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

Intuitively, this equation states that the decrease in τ_N that can occur after an increase in τ_W is proportional to the direct extra tax revenue induced by the increase in τ_W, plus a (typically negative) term representing any revenue gains that could occur from the behavioral effects of the tax change. 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 (22) 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 6. *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 $e^{z^\star D}_{\tau_K} = 1 + \frac{\mathbb{E}[\theta]}{\rho+\gamma}(1-\tau_K)\frac{\partial}{\partial \tau_K}\left(\tilde{r}^*_X \hat{k}(1) - \frac{(\phi(1-\epsilon)\hat{k}(1)\varphi)^2}{2}\right)$.

4 Optimal Taxes

In this section, we formulate and solve the planner's problem of choosing optimal taxes to maximize the steady-state welfare of a newborn agent. We use a perturbation approach, which requires that we first characterize the marginal effects of changes in tax rates on welfare, which are then all equal to zero at the optimum.

4.1 Effects of Tax Changes on Welfare

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, we derive here an expression for the welfare of a newborn worker in the steady state, $V^N(F^N, X)$. It is then relatively straightforward to extend the analysis to include the welfare of entrepreneurs.

Effect of A Tax Change on Worker Lifetime Utility We can derive a simple formula for the change in $V^N(F^N, X)$ from a marginal change in tax rates. Envelope theorem arguments, formalized in Appendix B.6, imply that this can be calculated as if the worker changes first period consumption by an amount equal to the change in lifetime resources from the tax reform, and keeps future consumption unchanged. That is

$$\mathrm{d}V^N(F^N, X) = u'(c^N_0)\left(\sum_{s=0}^{\infty}\left(\frac{1-\gamma}{1+\tilde{R}_F}\right)^s\right)\left(\mathrm{d}\tilde{w} + \mathrm{d}\tilde{R}_F \mathcal{A}^N\right),$$

where $\mathrm{d}\tilde{w}$ and $\mathrm{d}\tilde{R}_F$ are the change in $\tilde{w}$ and $\tilde{R}_F$ as a result of the tax change and where

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

is the average value of the worker's discounted lifetime assets.

Substituting in the worker's optimal consumption decision from Section 3.1, we obtain

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

Intuitively, the change in a newborn worker's expected lifetime utility is proportional to the change in worker income on the margin due to changes in $\tilde{w}$ and $\tilde{R}_F$. We extend this to the continuous time case in the following Lemma.

Lemma 4. *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 $\mathcal{A}^N = (\gamma + \tilde{R}_F)\int_{s=0}^{\infty} e^{-(\gamma+\tilde{R}_F)s}a_s^N ds$.

Proof. See Appendix B.6. □

Welfare Effect of A Tax Change To construct the measure of welfare $\mathcal{W}$, recall that

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

Then, it follows that the expected lifetime utility of a newborn agent is

$$\begin{aligned}\mathcal{W} &= \int_{-\infty}^{\infty} \max\left\{V^N(F^N, X); \mathbb{E}_\theta V(F, \theta, X) + \int_{s=0}^{\infty} e^{-(\rho+\gamma)} z\mathrm{d}s\right\}\mathrm{d}H_z(z)\\ &\equiv V^N(F^N, X) + \frac{1}{\rho+\gamma}\int_{z^\star}^{\infty}(z - z^\star)\mathrm{d}H_z(z).\end{aligned}$$

A first order approximation of this gives the change in welfare from a small tax reform

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

Combining with the occupational choice condition $N = H_z(z^\star)$ and using $e^N_{\tilde{w}} = \frac{H'(z^\star)}{N}$ gives

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

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

$$\tilde{w}\Delta = d\tilde{w} + \mathcal{A}^N \mathrm{d}\tilde{R}_F - \frac{(1-N)\,\tilde{w}\mathrm{d}N}{N e^N_{\tilde{w}}}.$$

The first two terms denote the change in the welfare of workers. The final term represents the fact that, if a tax reform increases N, then it must be making relatively entrepreneurs worse off, so the increase in aggregate welfare is less than the increase in worker welfare.

We can study the partial equilibrium effects of tax changes on welfare using the same approach as in Section 3.5. Consider first a change in τ_K, with a corresponding change in τ_N to balance the budget. The change in consumption equivalent welfare satisfies

$$\Delta = -\frac{\mathrm{d}\tau_N}{1-\tau_N} - \left(\frac{(r_F-\delta)\,\mathcal{A}^N}{(1-\tau_N)\,w}\right)\mathrm{d}\tau_K - \frac{(1-N)\,\mathrm{d}N}{N e^N_{\tilde{w}}}.$$

As in Section 3.5, we infer $\mathrm{d}\tau_N$ as a function of $\mathrm{d}\tau_K$ by differentiating the government's budget constraint, holding constant pre-tax prices. We obtain

$$0 = B_{\tau_N}\mathrm{d}\tau_N + B_{\tau_K}\mathrm{d}\tau_K + \sum_j \tau_j \frac{\partial B_{\tau_j}}{\partial \tau_K}\mathrm{d}\tau_K.$$

Substituting this into our expression for Δ above and rearranging, we obtain

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

where $B^N_{\tau_K} = (r_F-\delta)\,\mathcal{A}^N$ denotes the lifetime average additional tax payments a worker would have to make, all else equal, after a unit rise in τ_K. Intuitively, the change in a worker's welfare from an increase in τ_K is proportional to revenue gained (since this decreases in τ_N) minus the component of the τ_K tax rise that is paid for on average by workers over their lifetime. The revenue gain per unit rise in τ_K is the tax base B_{τ_K}, plus a (negative) term representing the loss in revenue that arises from the behavioral response.

The proposition below establishes a formula for the effect of any tax change on welfare.

Proposition 7. *The partial equilibrium effect of a small change in tax rate τ_j, $j \in \{K;W\}$, on welfare satisfies*

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

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

In the next section, we show that the optimal tax rates can be written as functions of the derivatives of the B_{τ_j} terms.

4.2 Optimal Tax Formula

Optimal taxes depend only on partial equilibrium effects of tax changes – general equilibrium effects via price changes can be ignored. This is because, as shown in Proposition 2, the steady state of the economy depends (locally) on post-tax prices not pre-tax prices. Therefore, the problem of choosing optimal tax rates can be recast as a problem of choosing optimal post-tax prices, which is equivalent to choosing optimal tax rates holding pre-tax prices fixed, as in, for instance Diamond and Mirrlees (1971) and Piketty and Saez (2013). Then, choosing τ_K and τ_W optimally simply means that, at the optimal tax rates, $\Delta = 0$ for any small tax change $\mathrm{d}\tau_j$. The optimal choice of τ_N can then be inferred from government budget balance. Specifically, the first order condition for each $\tau_j \in \{\tau_K, \tau_W\}$ is

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

Using the government budget constraint we can eliminate τ_N and rearrange to obtain

$$\begin{aligned} 0 \;&=\; (1 + A_j \tau_j)\left(B_{\tau_j} - N B^N_{\tau_j}\right) - \frac{(1-N) B_{\tau_N} e^N_{\tau_j}}{e^N_{\tilde{w}}} + \overline{G} e^N_{\tau_j} \left(1 + \frac{1-N}{e^N_{\tilde{w}}}\right) \\ &+ \sum_{m \in \{K;W\}} \tau_m B_{\tau_m} \left[e^{B_{\tau_m}}_{\tau_j} - e^N_{\tau_j} \left(1 + \frac{1-N}{e^N_{\tilde{w}}}\right)\right], \end{aligned}$$

where $A_{\tau_K} = -1$ and $A_{\tau_W} = 0$.

The first order conditions can be written in matrix form as follows

$$\begin{aligned} \mathbf{0} \;&=\; \left(B - B^N\right)\mathbf{1} - A\left(B - B^N\right)\mathcal{T} - (1-N)\left(e^N_{\tilde{w}}\right)^{-1} B_{\tau_N} \mathbf{e}^N + \overline{G}\left(1 + (1-N)\left(e^N_{\tilde{w}}\right)^{-1}\right)\mathbf{e}^N \\ &+ \mathcal{E} B \mathcal{T} - \mathbf{e}^N \mathbf{1}^T B \mathcal{T} \left(1 + (1-N)\left(e^N_{\tilde{w}}\right)^{-1}\right), \end{aligned}$$

where $\mathbf{1}$ denotes the column vector $(1,1)^T$, $\overline{\mathbf{e}}^N$ denotes the column vector $(\overline{e}^N_{\tau_K}, \overline{e}^N_{\tau_W})^T$, $\mathcal{T}$ denotes the column vector $(\tau_K, \tau_W)^T$, and A, B, B^N and $\mathcal{E}$ are defined as follows

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} B_{\tau_K} & 0 \\ 0 & B_{\tau_W} \end{pmatrix}, \quad B^N = \begin{pmatrix} N B^N_{\tau_K} & 0 \\ 0 & N B^N_{\tau_W} \end{pmatrix}, \quad \mathcal{E} = \begin{pmatrix} \overline{e}^{B_{\tau_K}}_{\tau_K} & \overline{e}^{B_{\tau_W}}_{\tau_K} \\ \overline{e}^{B_{\tau_K}}_{\tau_W} & \overline{e}^{B_{\tau_W}}_{\tau_W} \end{pmatrix}.$$

Rearranging, we obtain the vector of optimal taxes, summarized in the following proposition.

Proposition 8. *The optimal steady state tax vector $\mathcal{T} = [\tau_K; \tau_W]^T$ is given by*

$$\mathcal{T} = \left(A - g_1 + B^{-1}\left(-\mathcal{E} + \mathbf{e}^N \mathbf{1}^T\right) B\right)^{-1} \left(\mathbf{1} - \mathbf{g}_2 + B^{-1}\overline{G}\mathbf{e}^N\right), \tag{23}$$

where

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

Despite the complexity of the model, optimal taxes depend only on the size of the tax bases, and the elasticities of tax bases with respect to taxes. There are noticeable similarities to scalar optimal linear tax formulae in the literature. These formulae often take a form similar to $\tau = \frac{1-g}{1-g+|e|}$, 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. 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. The reason that the vector $\mathbf{e}^N$ appears in the formula is that if taxes on workers are higher than on entrepreneurs on average, then the tax revenue increases from shifting agents to becoming workers, thus the degree to which taxes affect occupational choice matters.

The matrix g_1 and 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 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 relative 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 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 and wealth would be zero. This signifies what the motives to tax capital income and wealth are: taxes on capital income and wealth are valuable for redistributive reasons (i.e. if $\mathbf{g}_2 \neq \mathbf{1}$) and also valuable to discourage excessive entry into entrepreneurship (indicated by $\mathbf{e}^N$).

4.3 Generality of the Optimal Tax Formula

In this section, we show that the derivation of the optimal tax formula obtained in Proposition 8 does not depend on many specific features of the model, including the financial friction, the functional form determining entrepreneurial risk, and the logarithmic utility.

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 8 are finite and non-zero. Then, optimal taxes are given by the formula in Proposition 8.

Proof. See Appendix B.7 □

The specific assumptions we made on the utility function, entrepreneurial risk and agency frictions are necessary for characterizing elasticities and for infering values for optimal taxes.

4.4 A Numerical Calibration

To interpret the magnitudes of these optimal taxes in practice, we undertake a numerical calibration of the economy outlined in Section 2, and set taxes at their current levels in the United States. We summarize parameter values in Table 1.

Assigned Parameters. We calibrate the economy at annual frequency.
Demographics and Technology 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. We set the distribution of z, H_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 (have $z_i > 0$). We set $\epsilon_E = 1.5$ as a baseline and explore below how results 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 the uniform distribution on $[0, 1]$ and the annual 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 $\phi = 0.76$ 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.[20] 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 (Mehrotra and Crouzet, 2017, Boar and Midrigan, 2019).

Initial Tax system We set τ_K to 20%, in line with the US corporate tax rate for small businesses reported in the OECD Tax Database, and τ_W to zero, in line with the current practice in the US. 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.

Calibrated Parameters. We assume that final output is produced with the technology $Y = Y_E^{\alpha_E} Y_F^{\alpha_F} N^{1-\alpha_E-\alpha_F}$ and calibrate α_E and α_F jointly with the remaining parameters $\rho, \underline{k}_E, \varphi, h_0$ and τ_N, 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, the fraction of households who enjoy entrepreneurship, and the labor tax rate. We set these seven 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 by entrepreneurs, the share of households who are entrepreneurs. As a seventh condition, we ensure that τ_N balances the government's budget. We assume a labor share of 2/3 and a capital output ratio of 3, roughly 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%.[21] 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 (2020) respectively using data from the Survey of Consumer Finances.

Wealth Distribution. In our calibration, we only targeted the wealth share of entrepreneurs, but the model matches top wealth inequality more broadly, as shown in Table 2, lending credibility to the model for studying the optimal capital and wealth taxation.

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

[21] 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.

Table 1: Parameter Values

Parameter	Value used	Target moment
γ	0.025	Working Life: 40 Years
δ	0.07	Depreciation
ϵ_E	1.5	Entry elasticity
λ_θ	0.115	Profitability autocor. (Cooper and Haltiwanger, 2006)
ϕ	0.840	Small Bus. Owner Equity Share (SSBF)
$\underline{\epsilon}$	0.350	Debt-to-asset ratio (Boar and Midrigan, 2019)
τ_K	0.200	Corp. tax rate small businesses (OECD Tax Database)
τ_W	0	Current US level
τ_N	0.263	Gov. budget balance
$\bar{G}$	0.200	Govt. spending/GDP
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

Table 2: Wealth Inequality

Wealth Share	Model	Data (Piketty et al., 2018)	Data (Smith et al., 2021)
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%

Implied Values of Elasticities. We first calculate the values of the tax bases and elasticities used in the optimal tax formula in Proposition 8. The calibrated values of the terms in the vectors $\mathbf{g}_2$, $\mathbf{e}^N$ and the matrix $\mathcal{E}$ are as follows[22]

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

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 government puts substantial weight on the welfare of capital

[22]The calibrated values of the remaining terms in the optimal tax formula are shown in Appendix B.8.

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 $\mathbf{e}^N$, containing elasticities slightly above zero, signifies that increases in capital and wealth taxes somewhat increase the number of agents who become workers, but not substantially. The elasticity matrix $\mathcal{E}$, however, contains very large negative elasticities. The diagonal terms indicate that a 1% decrease in post-tax capital income reduces the tax base by 5.2% and a 1% wealth tax on the margin reduces aggregate wealth by 293%. The latter number seems very 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. Leting $e^{B_{\tau_W}}_{-\tilde{R}_F} \equiv \tilde{R}_F e^{B_{\tau_W}}_{\tau_W}$ 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.[23] 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.

Optimal Taxes. We calculate optimal tax rates by imputting the initial steady state values of the tax bases and elasticities into the formula in Proposition 8. The results, reported in the first row of Table 3, indicate that it is optimal to tax both capital income and wealth to some degree, with capital income taxes significantly below current levels.

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. To obtain the exactly optimal tax rates, we must apply the optimal tax formula recursively by calculating (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. The second row of Table 3 reports

[23] We conjecture that the reason that our elasticity is still noticeably 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.

the exactly optimal taxes and shows that they are reasonably close to the approximate values. That being said, the exact optimal taxes do involve slightly lower capital income taxes and slightly higher wealth taxes. This is because cutting capital income taxes and raising wealth taxes increases the magnitude of the elasticity of tax bases with respect to the capital income tax compared to the initial steady state, encouraging the planner to move away from capital income taxes. An implication of Table 3 is that optimal tax rates on capital income, labor income and wealth are not hugely different from current US levels. Unsurprisingly, then, we find that the consumption equivalent welfare gains from shifting from status quo taxes to the optimal welfare-maximizing taxes are relatively negligible – 0.2% of current consumption.

To better understand the key factors determining these optimal tax rates, the last two rows of the table show the tax rates on capital and wealth that a planner would set if intending to maximize the revenue from these taxes. This is identical to applying the formula in Proposition 8 while setting the terms in g_1 and $\mathbf{g}_2$ to zero. 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 that can be calculated recursively. Optimal revenue-maximizing taxes are much higher on capital income, and essentially zero on wealth. This reveals that the main reason that it is 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. Since the g_1 and $\mathbf{g}_2$ terms are not involved in revenue-maximizing taxes, capital income is taxed much more heavily.

Table 3: Optimal Taxes Under Welfare and Revenue Maximization

	τ_K^*	τ_W^*	τ_N^*
Welfare-Maximization, Initial Elasticities	8.3%	0.1%	27.6%
Welfare-Maximization, Exact Opt. Taxes	3.7%	0.2%	28.0%
Revenue-Maximization, Initial Elasticities	22.1%	0.0%	25.9%
Revenue Maximization, Exact Opt. Taxes	20.4%	0.0%	26.2%

The Importance of Financial Frictions and Entrepreneurial Entry. To further understand the motivations behind capital income and wealth taxes in the model, we show how the exact optimal taxes vary 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 with respect to taxes in Figure 3. The top row of each figure shows the optimal welfare-maximizing 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_N). The blue line on the bottom row shows the elasticities at the initial steady state, and the red line shows the values of the elasticities at the optimal tax rates.

The figures reveal that, for low values of ϕ, $\underline{k}_E$ and ϵ_E, it is optimal to set negative capital income taxes and positive wealth taxes and that the elasticities of capital income and wealth with respect to taxes become very large. On the other hand, higher values of ϕ, $\underline{k}_E$ and ϵ_E yield very similar values of the optimal taxes and elasticities to the baseline calibration.

To understand the intuition behind these effects, recall from Section 3.5 that a determinant of the elasticity of K_E, and therefore capital income, with respect to taxes is $1-\frac{(1-N)\underline{k}_E}{K_E}$. Changes in ϕ and $\underline{k}_E$ mainly affect optimal taxes by affecting $\frac{(1-N)\underline{k}_E}{K_E}$, which represents the part of K_E which arises from each entrepreneur, inelastically, allocating $\underline{k}_E$ to the risky technology. As such, $\frac{(1-N)\underline{k}_E(r_E-r_F)}{K_E}$ represents the part of capital income which is relatively inelastic with respect to taxes as is due to economic profits earned by all entrepreneurs equally as a consequence of the fact that $r_E > r_F$, and does not represent the return to savings. Therefore, when $\frac{(1-N)\underline{k}_E}{K_E}$ is small 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 $\tilde{r}_X$ that determines the allocation capital to the risky technology, whereas wealth taxes do not.[24] When ϵ_E, $\underline{k}_E$ and ϕ are small, these effects of capital income taxes become the dominant considerations when comparing capital income taxes and wealth taxes. In that case, it is optimal to set a negative tax on capital income, which shifts capital to the risky technology and leads to increases in entrepreneurs' income and aggregate wealth. This increases the revenue earned from wealth taxes by enough to compensate for the fiscal costs of negative capital income taxes.

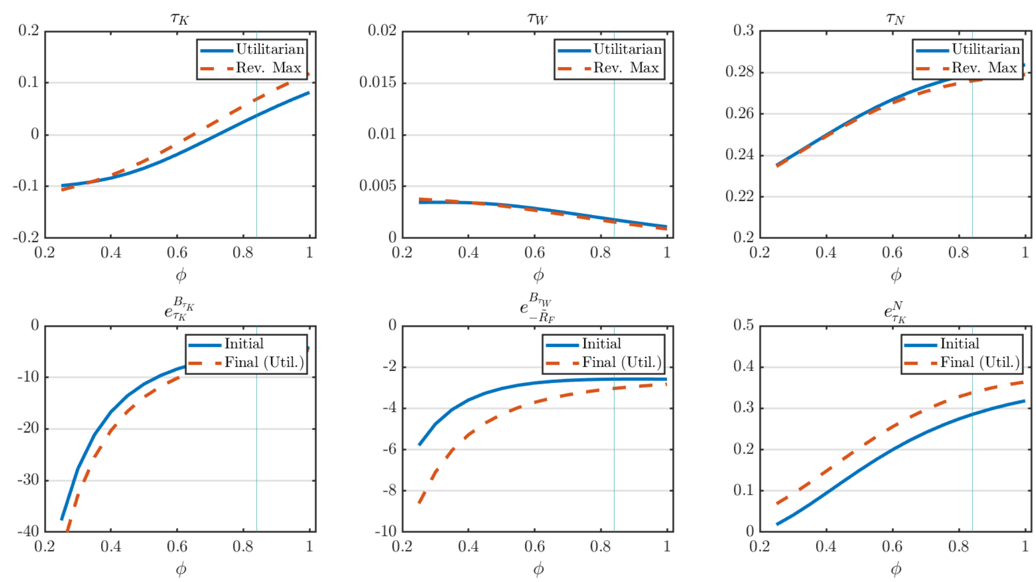
However, higher values of ϵ_E, $\underline{k}_E$ and ϕ cause counterveiling pressures that can motivate positive capital income tax rates. Higher $\underline{k}_E$ and ϕ both increase $\frac{(1-N)\underline{k}_E}{K_E}$, making capital income quite inelastic with respect to taxes. Since a large fraction of the capital income tax then falls on the inelastic $\underline{k}_E$ capital in the risky technology, capital income taxes are able to raise a lot more revenue relative to their effects on aggregate saving than wealth taxes. Equally, a high value of ϵ_E motivates a higher rate of capital income taxes because an additional disadvantage of low capital income taxes is that they encourage excessive 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.

In sum, this discussion suggests that the choice of optimal tax rates on capital income and wealth involve a tradeoff between four competing motivations. First, the effect of capital

[24] This is very similar to the 'use it or lose it' argument for taxing wealth in Guvenen et al. (2019).

income taxes on the allocation of capital motivates negative capital income taxes and positive wealth taxes. Second, some part of capital income is relatively insensitive with respect to taxes, which motivates higher taxes on capital income and lower taxes on wealth if $\frac{(1-N)\underline{k}_E}{K_E}$ is larger. Third, an additional motivation to tax capital income rather than wealth is to discourage excessive entry into entrepreneurship. Fourth, a redistribution motive favors taxing capital income less and wealth more, since capital income taxes partly fall on relatively poor entrepreneurs with a high marginal utility of consumption.

Figure 1: Optimal Taxes As Financial Frictions Vary

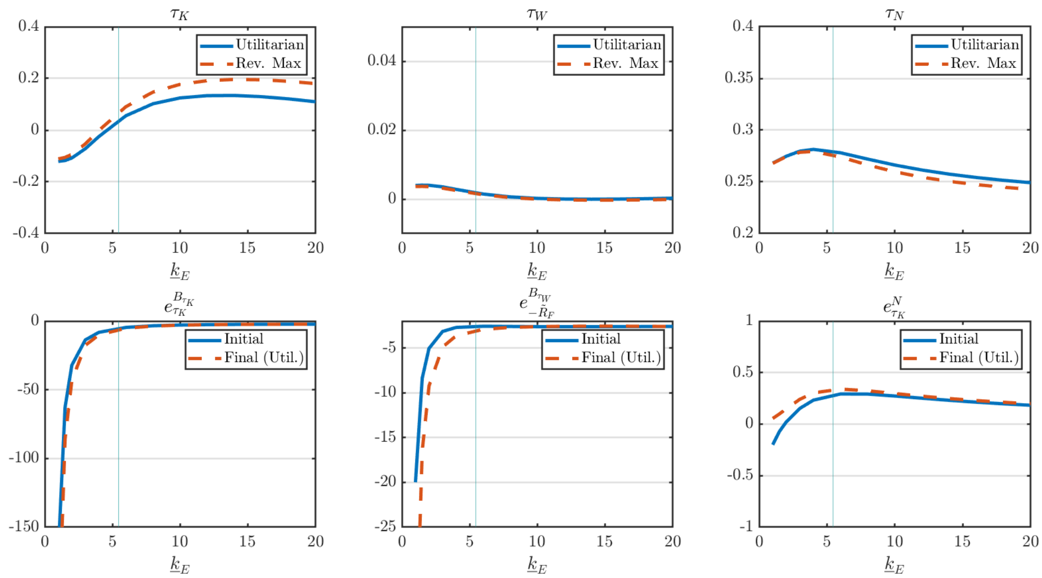


Notes: The vertical like denotes the benchmark value of ϕ.

Sensitivity analysis. Lastly, we explore how our results depend on other features of the environment. First, we replace our assumption of an endogenous financial frictions with an exogenous financial friction that is unaffected by taxes. To that end, we assume that $\hat{k}_E(\theta) = \overline{k}_E\theta$, where $\overline{k}_E$ is an exogenous constant unaffected by taxes. We calibrate $\overline{k}_E$, jointly with other parameters, so that the risky rate of return to capital (net of depreciation) is 8%, just as we calibrated φ in the baseline case. As discussed in Section 4.3, our theoretical results still go through in this case. We find, in this case, that the exact optimal taxes become $\tau_K = 8.0\%$, $\tau_W = 0.1\%$, and $\tau_N = 27.6\%$, so it is optimal to tax capital relatively more and to tax wealth and labor relatively less than in the baseline case. Intuitively, this is because the exogenous financial friction decreases the magnitude of $e^{\hat{k}_E}_{\tau_K}$ which reduces the elasticity of capital income with respect to taxes and makes capital income taxes more desirable.

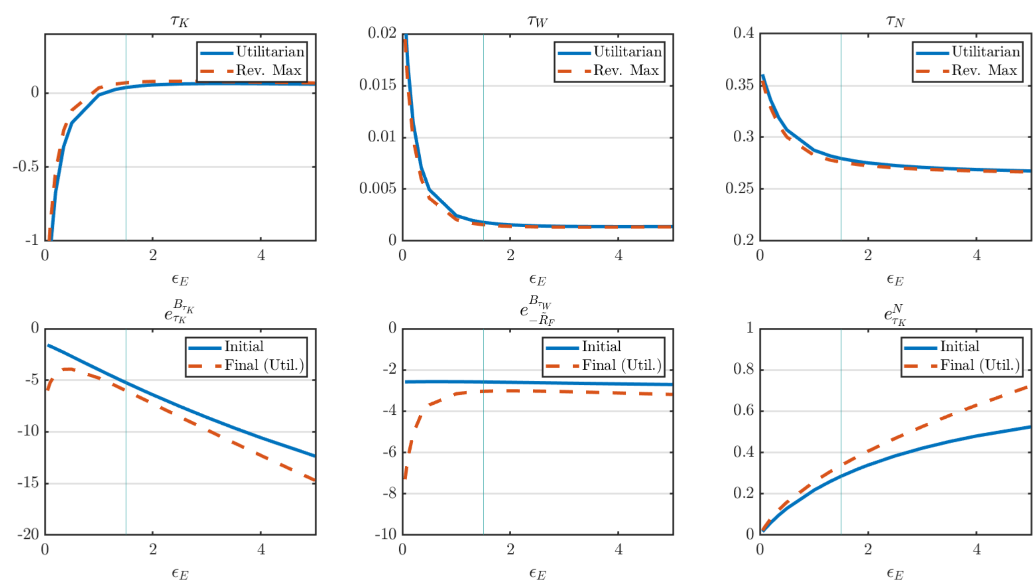
Second, we as is well known in the public finance literature, optimal taxes typically

Figure 2: Optimal Taxes As Returns to Scale Vary



Notes: The vertical like denotes the benchmark value of $\underline{k}_E$.

Figure 3: Optimal Taxes As Entry Elasticity Varies



Notes: The vertical like denotes the benchmark value of ϵ_E.

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.

5 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. 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 a redistributive 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. Nevertheless, the steady state welfare gains from moving from the current US tax policy to the optimal one are small. Our optimal tax formula is general and does not depend on many of the details of the model, such as functional form assumptions or the specific details of the financial friction.

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. (2019). Optimal positive capital taxes at interior steady states. Working Paper 25895, National Bureau of Economic Research.

Bernanke, B. S., Gertler, M., and Gilchrist, S. (1999). The financial accelerator in a quantitative business cycle framework. *Handbook of macroeconomics*, 1:1341–1393.

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.

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.

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.

Chetty, R. (2008). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. Working Paper 14399, National Bureau of Economic Research.

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

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.

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.

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

Guvenen, F., Kambourov, G., Kuruscu, B., Ocampo-Diaz, S., and Chen, D. (2019). Use it or lose it: Efficiency gains from wealth taxation. Technical report, National Bureau of Economic Research.

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

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.

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

Mehrotra, N. and Crouzet, N. (2017). Small and large firms over the business cycle. Working paper.

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 Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy*, 6(1):230–71.

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.

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.

Varian, H. R. (2014). *Intermediate microeconomics with calculus: a modern approach*. WW Norton & Company.

Appendices

A Discrete Time Model

A.1 Worker's Optimization Problem

We derive the solution to the worker's problem. We first rewrite the worker's problem, replacing a^N with P^N. Note that the worker's budget constraint may be rewritten as

$$c_t^N + R_{F,t}\frac{(1-\gamma)\,P_{t+1}^N}{R_{F,t}} = R_{F,t}P_t^N.$$

Hence, abusing notation, the worker problem can be written 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

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

Taking first order conditions, we get

$$\left(R_F P^N - (1-\gamma)P^{N'}\right)^{-1}(1-\gamma) = (1-\rho)(1-\gamma)\frac{\partial V^N(P^{N'}, X')}{\partial P^{N'}}.$$

Using the guess and simplifying, we obtain the worker's policy functions

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

To verify the guess, we plug the policy functions back into the value function

$$\begin{aligned} V^N(P^N, X) &= \max\log\left([1-(1-\rho)\,(1-\gamma)]\,R_F P^N\right) \\ &+ (1-\rho)\,(1-\gamma)\left(Q' + \frac{1}{1-(1-\rho)(1-\gamma)}\log\left((1-\rho)\,(1-\gamma)\,\frac{R_F}{1-\gamma}P^N\right)\right) \end{aligned}$$

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 Proof of Lemma 1

Consider an entrepreneur Alice, who at some period t has $s > 0$ times as much lifetime resources P as another entrepreneur Bob, and the same ability θ. The linearity of the constraints (2)-(9) implies that, in each period and each realization of past shocks, Alice could choose to set $k_E - \underline{k}_E$, k_F, k_H and $\hat{b} - R_F b$ at exactly s times the level that Bob would set given the same shock realizations. Then, the linearity of the $q(\cdot)$ function in k_E implies that, for each shock realization, Alice's $\tilde{k}_E - k_E$ would equal s times the value obtained by Bob under the same shock realization. Then, (4) and (7) imply that Alice could set c and P' at exactly s times the value set by Bob under the same shock realization. Since, for any c, $\log(sc) \equiv \log(s!`) + \log(c)$, Alice's present discounted utility from these choices would then be the same as Bob's plus an additional $\sum_{j=0}^{\infty}(1-\rho)^j(1-\gamma)^j \log(s) = \frac{\log(s)}{1-(1-\rho)(1-\gamma)}$.

Since these choices are possible for Alice, it must be that $V(sP, \theta, X) \geq \frac{\log(s)}{1-(1-\rho)(1-\gamma)} + V(P, \theta, X)$. However, on the other hand, Bob could equally choose to do everything that Alice does only $\frac{1}{s}$ times as much (i.e. setting $k_E - \underline{k}_E$ at $\frac{1}{s}$ of the value chosen by Alice etc.).. By the same logic as before, doing so would yield Bob a present discounted utility equal to Alice's minus $\frac{\log(s)}{1-(1-\rho)(1-\gamma)}$. Therefore, it must be the case that $V(P, \theta, X) \geq V(sP, \theta, X) - \frac{\log(s)}{1-(1-\rho)(1-\gamma)}$. Comparing these two inequalities that V must fulfil, it is immediate that it cannot satisfy both unless $V(sP, \theta, X) = \frac{\log(s)}{1-(1-\rho)(1-\gamma)} + V(P, \theta, X)$. In that case, it must be that $V(P, \theta, X) \equiv \frac{\log(P)}{1-(1-\rho)(1-\gamma)} + V(1, \theta, X)$. Let $\overline{V}(\theta, X)$ denote $V(1, \theta, X)$. Then it follows that $V(P, \theta, X) = \overline{V}(\theta, X) + \frac{\log(P)}{1-(1-\rho)(1-\gamma)}$.

Using Lemma 1, the solution to the between period problem can be found by taking the first order condition and combining it with equation (4) to conclude that

$$c = (1 - (1-\rho)(1-\gamma))\,\omega \qquad \text{and} \qquad P' = (1-\rho)\,\omega.$$

Substituting these choices into the Bellman equation (3), we have that

$$\begin{aligned}\tilde{V}(\omega, X) &= \frac{\log(\omega)}{1-(1-\rho)(1-\gamma)} + \log(1-(1-\rho)(1-\gamma)) \\ &+ \frac{(1-\rho)(1-\gamma)\log((1-\rho))}{1-(1-\rho)(1-\gamma)} + (1-\rho)(1-\gamma)\,\mathbb{E}\left[\overline{V}(\theta', X')\right].\end{aligned}$$

□

A.3 Proof of Lemma 2

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. We show that, for any choices of k_E and k_F by the entreprenueur, this contract yields the same value of ω in every state of the world as the equilibrium contract, so that the two contracts are equivalent.

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

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

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 (7) 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 (7) implies that

$$V_P = R_F P + R_F b + [(1-\tau_K)\, r_E + 1 - \delta](\tilde{k}_E - \underline{k}_E) + (\tau_K \delta - \tau_W - R_F)\,(k_E - \underline{k}_E). \tag{A.2}$$

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

$$b = \tilde{b} + \frac{s(E_\xi[V_P] - R_F \tilde{b})}{R_F} \tag{A.3}$$

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

$$\omega = (1-s)(V_P - R_F\tilde{b}). \tag{A.4}$$

Substituting into this equations (A.1), (A.2) and (A.3) 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 the no arbitrage condition $R_{F,t} = 1 + (1 - \tau_{K,t})(r_{F,t} - \delta) - \tau_{W,t}$, this is

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

This is the same expression for ω as in the equilibrium financial contract. So the equilibrium financial contract delivers the same ω in every state of the world as a contract in which the entrepreneur sells equity and issues risk-free debt. For the two contracts to be equivalent, the level of debt $\tilde{b}$ issued by the entrepreneur in the case with equity and debt is uniquely determined by (A.3) and (A.4), 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.4) 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

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

We let $P_{i,t}$ denote the the lifetime income of an entrepreneur who puts exactly $\underline{k}_E$ capital into the risky technology, no capital into the risk-free technology and lends her remaining endowment to the bank at risk free rate $1 + (R_F - 1)\Delta = 1 + \tilde{R}_F\Delta$. It follows that,

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

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

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

First, consider the worker's problem for a given period length Δ

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

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, $a^{N\prime}$.

Second, 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

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

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:

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

The equilibrium conditions of the model with period length Δ are summarized below.

Definition 4. *Given a sequence of tax rates* $\{\tau_{W,t}, \tau_{K,t}, \tau_{N,t}\}_{t=0}^{\infty}$, *an* ***equilibrium*** $\mathcal{E}_\Delta$ *of the economy with period length* Δ *is a sequence of prices* $\{R_{F,t}, r_{E,t}, r_{F,t}, w_t\}_{t=0}^{\infty}$, *policy functions giving entrepreneurs' and workers' decisions and a sequence of aggregate variables* $\{C_t, K_t, K_{E,t}, K_{F,t}, K_{H,t}, Y_t, N_t\}_{t=0}^{\infty}$ *such that:*

1. *The government's budget constraint is balanced every period:*

$$\overline{G}\Delta = \tau_{N,t} w_t \Delta N_t + \tau_{K,t}(Y_t \Delta - w_t \Delta N_t - \delta \Delta K_t) + \tau_{W,t} \Delta K_t. \tag{B.6}$$

2. *Workers' decision rules solve the worker's optimization problem B.3.*

3. *Entrepreneurs' decision rules are the solution to the entrepreneur's problem B.4.*

4. $\{C_t, K_t, K_{E,t}, K_{F,t}, K_{H,t}, Y_t\}_{t=0}^{\infty}$ *represent the aggregate of household's decisions.*

5. *The indifference value* z_t^* *for agents' occupational choice is given by* (B.4)

6. *The aggregate level of* N_t *evolves according to* (B.5)

7. *The asset, intermediate goods and labor markets clear.*

The final goods market clearing condition then follows by Walras' law

$$\overline{G}_t \Delta + C_t \Delta + K_{t+\Delta} \quad = \quad Y_t \Delta + (1 - \delta\Delta)\left(K_{E,t} + K_{F,t} - K_{H,t}\right) + C_{H,t}\Delta, \tag{B.7}$$

where $C_{H,t}\Delta = \phi K_{H,t}$.

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

Following the same derivation as in the main text, where $\Delta = 1$, it can readily be shown that the solution to the worker's problem is given by

$$c^N \Delta = [1-(1-\rho\Delta)(1-\gamma\Delta)]\left(1+\tilde{R}_F\Delta\right)P^N = (\gamma+\rho-\gamma\rho\Delta)\left(1+\tilde{R}_F\Delta\right)\Delta P^N, \quad \text{(B.8)}$$

$$P^{N\prime} = (1-\rho\Delta)(1+\tilde{R}_F\Delta)P^N, \quad \text{(B.9)}$$

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

$$c\Delta = (1-(1-\rho\Delta)(1-\gamma\Delta))\omega = (\gamma+\rho+\gamma\rho\Delta)\Delta\omega, \quad \text{(B.10)}$$

$$P' = (1-\rho\Delta)\omega = \frac{1}{1-\gamma\Delta}(\omega - c\Delta). \quad \text{(B.11)}$$

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

$$\sup \int_{\xi} \log\left(\underline{\omega} + \phi(\tilde{k}_E - \underline{k}_E)\right) \mathrm{d}H_{\xi}(\xi), \quad \text{(B.12)}$$

subject to the constraints:

$$\underline{\omega} = (-\phi + (r_E - r_F)(1-\tau_K)\Delta)(k_E - \underline{k}_E) + (1+\tilde{R}_F\Delta)P \quad \text{(B.13)}$$

$$k_E \geq \underline{k}_E \quad \text{(B.14)}$$

$$0 \leq \underline{\omega} + \phi\underline{\epsilon}(k_E - \underline{k}_E), \quad \text{(B.15)}$$

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

Proposition 10. *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:*

$$k_E(P,\theta,X) - \underline{k}_E = \frac{S_\theta^{-1}\left(\max\left\{0;\min\left\{\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi};S_\theta^\star\right\}\right\}\right)P\left(1+\tilde{R}_F\Delta\right)}{\phi-(r_E-r_F)\Delta(1-\tau_K)S_\theta^{-1}\left(\max\left\{0;\min\left\{\frac{(r_E-r_F)\Delta(1-\tau_K)}{\phi};S_\theta^\star\right\}\right\}\right)}, \quad \text{(B.16)}$$

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

$$\omega = (\phi(\epsilon - 1) + (1-\tau_K)(r_E - r_F)\Delta)(k_E - \underline{k}_E) + \left(1 + \tilde{R}_F \Delta\right) P \tag{B.17}$$

$$c = (\gamma + \rho + \gamma\rho\Delta)\omega \tag{B.18}$$

$$P' = (1-\rho\Delta)\omega, \tag{B.19}$$

where

$$\epsilon = 1 + (1 - \underline{\epsilon})\left(\exp\left(\frac{\varphi\xi_{i,t}\sqrt{\Delta}}{\sqrt{\theta_{i,t}}} - \frac{\varphi^2\Delta}{2\theta_{i,t}}\right) - 1\right), \tag{B.20}$$

$$S_\theta(x) = 1 - \frac{\int_\xi \left(1 + x(\epsilon - 1)\right)^{-1} \epsilon H_\xi(\xi)\, d\epsilon}{\int_\xi \left(1 + x(\epsilon-1)\right)^{-1} H_\xi(\xi)\, d\epsilon}, \quad \forall x \in \left[0, \frac{1}{1-\underline{\epsilon}}\right], \tag{B.21}$$

and

$$S_\theta^\star = S_\theta\left(\frac{1}{1-\underline{\epsilon}}\right). \tag{B.22}$$

Here $S_\theta : \left[0, \frac{1}{1-\underline{\epsilon}}\right] \to [0, S_\theta^\star]$ *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

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

where

$$x = \frac{\phi(k_E - \underline{k}_E)}{\underline{\omega} + \phi(k_E - \underline{k}_E)}. \tag{B.23}$$

Here we used that $\underline{\omega} + \phi(k_E - \underline{k}_E) > 0$ at any feasible k_E. This holds because the constraint (B.15) implies that

$$0 \le \underline{\omega} + \phi\underline{\epsilon}(k_E - \underline{k}_E) < \underline{\omega} + \phi\mathbb{E}[\epsilon](k_E - \underline{k}_E) = \underline{\omega} + \phi(k_E - \underline{k}_E).$$

Since $k_E - \underline{k}_E \ge 0$, equation (B.23) in turn implies that $x \ge 0$ at any feasible choice. Furthermore, $x \le \frac{1}{1-\underline{\epsilon}}$. To show this, note that it was shown above that the entrepreneur's end of period consumption is proportional to

$$\omega \equiv \underline{\omega} + \phi\epsilon(k_E - \underline{k}_E) \equiv (1 - x + x\epsilon)(\underline{\omega} + \phi(k_E - \underline{k}_E))$$

Since $\underline{\omega} + \phi(k_E - \underline{k}_E) > 0$, it follows that $1 - x + x\underline{\epsilon} \geq 0$, or the entrepreneur's consumption could be negative. Rearranging this condition, we obtain that $x \leq \frac{1}{1-\underline{\epsilon}}$, as desired.

Now we show that x is monotonically increasing in k_E. Using the definition of $\underline{\omega}$ in equation (B.15) and the definition of x in equation (B.23), we obtain that

$$\frac{\partial x}{\partial k_E} \propto \left(1 + \tilde{R}_F \Delta\right) P.$$

Since the entrepreneur can convert units of capital into consumption at rate ϕ and the risk free rate of return is $(1+\tilde{R}_F\Delta)$, there can only be an equilibrium in which some entrepreneurs put a positive amount of capital in the risk-free sector if $(1 + \tilde{R}_F\Delta) > 0$. Moreover, by definition $P > 0$, leading us to conclude that $\frac{\partial x}{\partial k_E} > 0$. So x is monotonically increasing in k_E, and $x = 0$ when $(k_E - \underline{k}_E) = 0$, while $x = \frac{1}{1-\underline{\epsilon}}$ corresponds to the highest possible k_E the entrepreneur can choose while ensuring that consumption is non-negative.

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

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

where

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

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

$$\frac{(r_E - r_F)\Delta(1 - \tau_K)}{\phi} - S_\theta(x). \tag{B.24}$$

In the case of an interior solution, the first order condition for the optimal choice of k_E is

$$\frac{(r_E - r_F)\Delta(1-\tau_K)}{\phi} - S_\theta(x) = 0. \tag{B.25}$$

We now show that $S_\theta(x)$ is strictly increasing for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$. We rewrite $S_\theta(x)$ as

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

Since all the terms in $S_\theta(x)$ are differentiable with respect to x, for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, it follows that $S_\theta(x)$ is itself differentiable over $x \in (0, \frac{1}{1-\underline{\epsilon}}]$. The derivative is

$$S'_\theta(x) = x^{-2}\left(-1 + \left(\int_\xi \Big(x^{-1} + \epsilon - 1\Big)^{-1} \mathrm{d}H_\xi(\xi)\right)^{-2} \int_\xi \Big(x^{-1} + \epsilon - 1\Big)^{-2} \mathrm{d}H_\xi(\xi)\right). \tag{B.26}$$

Using Jensen's inequality,

$$\mathbb{E}\left[\Big(x^{-1} + \epsilon - 1\Big)^{-2}\right] = \mathbb{E}\left[\left(\Big(x^{-1} + \epsilon - 1\Big)^{-1}\right)^2\right] > \mathbb{E}\left[\Big(x^{-1} + \epsilon - 1\Big)^{-1}\right]^2.$$

Substituting this into equation (B.26)

$$S'_\theta(x) > x^{-2}\left(-1 + \left(\int_\xi \Big(x^{-1} + \epsilon - 1\Big)^{-1} \mathrm{d}H_\xi(\xi)\right)^{-2} \left(\int_\xi \Big(x^{-1} + \epsilon - 1\Big)^{-1} \mathrm{d}H_\xi(\xi)\right)^2\right) = 0,$$

where we used that $x^{-1}+\epsilon-1 \geq 0$ for $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, with strict inequality for $\epsilon > 0$. Therefore, $S_\theta(x)$ is strictly increasing over $x \in (0, \frac{1}{1-\underline{\epsilon}}]$, as desired. Equation (B.21) 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 $x \in [0, \frac{1}{1-\underline{\epsilon}}]$.

Recall that the sign of the derivative of the entrepreneur's within-period objective function with respect to k_E is given by (B.24) and that x is monotonically increasing in k_E, with $x = 0$ when $(k_E - \underline{k}_E) = 0$ and $x = \frac{1}{1-\underline{\epsilon}}$ corresponding to the highest possible k_E the entrepreneur can choose. Since $S_\theta(x)$ is strictly increasing over $x \in [0, \frac{1}{1-\underline{\epsilon}}]$, it 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 (B.25). In that case, since the entrepreneur's within-period objective function is strictly concave with respect to k_E, it follows that the first order condition (B.25) characterizes the unique optimal choice of k_E.

For the case where the first order condition holds, we may use the fact that S is monotone and differentiable (and therefore invertible) to rearrange the first order condition as follows

$$x = S_\theta^{-1}\left(\frac{(r_E - r_F)\Delta(1-\tau_K)}{\phi}\right)$$

Let $S_\theta^\star = S_\theta\left(\frac{1}{1-\epsilon}\right)$. Then, $S_\theta^{-1}(S_\theta^\star) = \frac{1}{1-\epsilon}$. Furthermore, $S_\theta(0) = 0$ (by equation (B.21)) and so $S_\theta^{-1}(0) = 0$. Since S is monotonically increasing on $x[0, \frac{1}{1-\epsilon}]$, S_θ^{-1} is monotonically increasing on $[0, S_\theta^\star]$. As such, we can group the three cases above 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.23) 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 (13) and (B.13), 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.10) and (B.11) yields all the results of the proposition. □

B.3 Proof of Proposition 1

This proof makes use of the following two lemmas.

Lemma 5. *The following holds, for any* $x \in \left[0; \frac{1}{1-\epsilon}\right]$: $\lim_{\Delta\to 0} \frac{S_\theta(x)}{\Delta} = \frac{(1-\epsilon)^2\varphi^2 x}{\theta}$.

Proof. To prove this, note first that

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

Then, it remains to show that, for any x in the domain of S,

$$\lim_{\Delta\to 0}\left[\frac{\int_\xi \left(\frac{\epsilon-1}{\Delta}\right)(1+x(\epsilon-1))^{-1}\mathrm{d}H_\xi(\xi)}{\int_\xi (1+x(\epsilon-1))^{-1}\mathrm{d}H_\xi(\xi)}\right] = -\frac{x(1-\underline{\epsilon})^2\varphi^2\Delta}{\theta}. \tag{B.27}$$

To prove this, we prove the following results, from which equation (B.27) follows trivially

$$\lim_{\Delta\to 0}\int_\xi \left(\frac{\epsilon-1}{\Delta}\right)(1+x(\epsilon-1))^{-1}\mathrm{d}H_\xi(\xi) = -\frac{x(1-\underline{\epsilon})^2\varphi^2\Delta}{\theta} \tag{B.28}$$

$$\lim_{\Delta\to 0}\int_\xi (1+x(\epsilon-1))^{-1}\mathrm{d}H_\xi(\xi) = 1 \tag{B.29}$$

Using (B.20), we consider a first order approximation of $\epsilon - 1$ around $\sqrt{\Delta} = 0$

$$\epsilon - 1 = (1-\underline{\epsilon})\left(\exp\left(\frac{\varphi\xi\sqrt{\Delta}}{\sqrt{\theta}} - \frac{\varphi^2\Delta}{2\theta}\right) - 1\right) \simeq (1-\underline{\epsilon})\frac{\varphi\xi\sqrt{\Delta}}{\sqrt{\theta}}. \tag{B.30}$$

Similarly, in the neighborhood of $\sqrt{\Delta} = 0$

$$(1+x(\epsilon-1))^{-1} = \left(1 + x(1-\underline{\epsilon})\left(\exp\left(\frac{\varphi\xi\sqrt{\Delta}}{\sqrt{\theta}} - \frac{\varphi^2\Delta}{2\theta}\right) - 1\right)\right)^{-1} \simeq 1 - \frac{x(1-\underline{\epsilon})\varphi\xi\sqrt{\Delta}}{\sqrt{\theta}}$$

Multiplying the term in the limit on the left hand side of (B.28) by Δ, and ignoring terms of order greater than Δ, we therefore can write it as

$$\int_{-\infty}^{\infty}(1-\underline{\epsilon})\varphi\sqrt{\Delta}\frac{\xi}{\sqrt{\theta}}\left(1 - x(1-\underline{\epsilon})\varphi\sqrt{\Delta}\frac{\xi}{\sqrt{\theta}}\right)\mathrm{d}H_\xi(\xi) = \frac{(1-\underline{\epsilon})\varphi\sqrt{\Delta}E[\xi]}{\sqrt{\theta}} - \frac{x(1-\underline{\epsilon})^2\varphi^2\Delta E[\xi^2]}{\theta},$$

which is equal to $-\frac{x(1-\underline{\epsilon})^2\varphi^2\Delta}{\theta}$ since ξ has mean zero and variance 1. Hence (B.28) follows.

Likewise, considering the term on the left hand side of (B.29), we have

$$\int_{-\infty}^{\infty} 1 - \frac{x(1-\underline{\epsilon})\varphi\xi\sqrt{\Delta}}{\sqrt{\theta}}\mathrm{d}H_\xi(\xi) = 1 - \frac{x(1-\underline{\epsilon})\varphi E[\xi]\sqrt{\Delta}}{\sqrt{\theta}} = 1.$$

□

Lemma 6. *For any $z \in (-\infty, \infty)$, it holds that*

$$\lim_{\Delta\to 0} S_\theta^{-1}(\max\{0; \min\{z\Delta; S_\theta^\star\}\}) = \max\left\{0; \min\left\{\frac{z\theta}{(1-\underline{\epsilon})^2\varphi^2}; \frac{1}{1-\underline{\epsilon}}\right\}\right\}. \tag{B.31}$$

Proof. To prove (B.31), we first note that

$$\lim_{\Delta\to 0}\frac{S_\theta^\star}{\Delta}=\frac{(1-\underline{\epsilon})\varphi^2}{\theta}. \tag{B.32}$$

This follows immediately from the definition of $S_\theta^\star$ in (B.22) and from Lemma 5.

Now, we show that

$$\forall z\in\left[0,\frac{(1-\underline{\epsilon})\varphi^2}{\theta}\right),\quad \lim_{\Delta\to 0}S^{-1}(z\Delta)=\frac{z\theta}{\varphi^2(1-\underline{\epsilon})^2}. \tag{B.33}$$

To prove (B.33) , define the function $\mathscr{F}(x)$ according to

$$\mathscr{F}(x)=\frac{S_\theta(x)}{\Delta}. \tag{B.34}$$

where $x\in\left(0,\frac{1}{1-\underline{\epsilon}}\right)$. Given that S is continuous and strictly increasing, it follows that $\mathscr{F}(\cdot)$ is continuous and strictly increasing, and therefore invertible. We now show that, for any z in the range of $\mathscr{F}$,

$$\mathscr{F}^{-1}(z)\equiv S^{-1}(z\Delta). \tag{B.35}$$

Let $x=\mathscr{F}^{-1}(z)$. Then $z=\mathscr{F}(x)=\frac{S_\theta(x)}{\Delta}$, so $S_\theta(x)=z\Delta$ and $x=S_\theta^{-1}(z\Delta)$, giving (B.35).

Define the function $\overline{\mathscr{F}}(x)=(1-\underline{\epsilon})^2\varphi^2x$. Its inverse is

$$\overline{\mathscr{F}}^{-1}(x)=\frac{x\theta}{(1-\underline{\epsilon})^2\varphi^2}. \tag{B.36}$$

We know from Lemma 5 that, as $\Delta\to 0$, $\mathscr{F}(x)$ converges to $\overline{\mathscr{F}}(x)$. Since $\mathscr{F}$ and $\overline{\mathscr{F}}$ are continuous, this convergence is uniform, and the inverse $\mathscr{F}^{-1}(x)$ converges to $\overline{\mathscr{F}}^{-1}(x)$. Then, using equations (B.35) and (B.36), we obtain (B.33) for values of z in the relevant domain.

Note that, for $\Delta>0$, the domain of $S_\theta^{-1}(\cdot)$ is $[0,S_\theta^\star]$. Therefore, the result (B.33) must hold for all $z\in\left(0,\lim_{\Delta\to 0}\frac{S_\theta^\star}{\Delta}\right)\equiv\left(0,\frac{(1-\underline{\epsilon})\varphi^2}{\theta}\right)$ since, for any such z, $z\Delta$ in will be in the domain of S^{-1} for sufficiently small $\Delta>0$. Equally, it must be true that $\lim_{\Delta\to 0}S_\theta^{-1}(0)=0$, since $S_\theta^{-1}(0)=0$ for any value of $\Delta>0$. Therefore, (B.33) follows for $z\in\left[0,\frac{(1-\underline{\epsilon})\varphi^2}{\theta}\right)$.

Now, using (B.32) and (B.33), we prove (B.31) for $z\in(-\infty,\infty)$. We proceed in cases. First, consider the case $z\le 0$. Then, for sufficiently small $\Delta>0$, equation (B.32) implies that $z\Delta\le 0<S_\theta^\star$. Then,

$$\lim_{\Delta\to 0}S_\theta^{-1}(\max\{0;\min\{z\Delta;S_\theta^\star\}\})=\lim_{\Delta\to 0}S_\theta^{-1}(0)=0, \tag{B.37}$$

where the second equality used (B.33).

Now, suppose that $z\in\left(0,\frac{1}{1-\underline{\epsilon}}\right)$. Then, for sufficiently small $\Delta>0$, equation (B.32) implies that $0 < z\Delta < S_\theta^\star$. Then,

$$\lim_{\Delta \to 0} S_\theta^{-1}(\max\{0; \min\{z\Delta; S_\theta^\star\}\}) = \lim_{\Delta \to 0} S_\theta^{-1}(z\Delta) = \frac{z\theta}{(1-\underline{\epsilon})^2 \varphi^2}, \tag{B.38}$$

where the second equality used (B.33).

Now, suppose that $z > \frac{1}{1-\underline{\epsilon}}$. Then, for sufficiently small $\Delta > 0$, equation (B.32) implies that $z\Delta > S_\theta^\star > 0$. Then,

$$\lim_{\Delta \to 0} S_\theta^{-1}(\max\{0; \min\{z\Delta; S_\theta^\star\}\}) = \lim_{\Delta \to 0} S_\theta^{-1}(S_\theta^\star) = \lim_{\Delta \to 0} \frac{1}{1-\underline{\epsilon}} = \frac{1}{1-\underline{\epsilon}}, \tag{B.39}$$

where the second equality used that, for any $\Delta > 0$, $S_\theta^{-1}(S_\theta^\star) = \frac{1}{1-\underline{\epsilon}}$, since $S_\theta^\star = S\left(\frac{1}{1-\underline{\epsilon}}\right)$.

Comparing equation (B.31) with (B.37),(B.38) and (B.39), we see that we have proven (B.31) for any $z \in (-\infty, \infty)$ except for $z = \frac{1}{1-\underline{\epsilon}}$. In particular, (B.31) must hold for all $z \neq \frac{1}{1-\underline{\epsilon}}$ in the neighborhood of $\frac{1}{1-\underline{\epsilon}}$. Note that the left hand side of (B.31) is continuous and weakly increasing in z for any $\Delta > 0$. The right hand side of (B.31) is continuous and weakly increasing in z. Continuity arguments then imply that (B.31) also holds at $z = \frac{1}{1-\underline{\epsilon}}$. □

Proof of the Propositon

First, we derive the worker's continuous time solution. Consider the discrete time solution with period length Δ

$$c^N \Delta = (\gamma + \rho - \gamma\rho\Delta)\,(1 + \tilde{R}_F \Delta)\Delta P^N, \tag{B.40}$$

$$P^{N'} = (1 - \rho\Delta)\,(1 + \tilde{R}_F \Delta) P^N. \tag{B.41}$$

Taking the limit of equation (B.40) 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.16) as $\Delta \to 0$, and using Lemma 6, implires (14). Combining (B.17) and (B.18) 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 (14). Finally, we note that Proposition 10 implies that

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

so that

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

Considering Δ close to zero, using (B.30), and dropping higher order terms in Δ, we obtain

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

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.

It remains to show that in the continuous time case, all entrepreneurs choose $k_H = 0$. The entrepreneur will not hide capital if $\phi < (1 - \tau_{K,t})r_{E,t} + (1 - \delta)$. With period length Δ, this inequality immediately becomes $\phi < (1 - \tau_{K,t})r_{E,t}\Delta + 1 - \delta\Delta$, In the limit as $\Delta \to 0$, this is $\phi < 1$, which is trivially satisfied since $\phi \in (0, 1)$ by assumption.

□

B.4 Proof of Proposition 2

As a first step, we characterize a steady state equilibrium in Proposition 11, as per Definitions 1 and 2. As a second step, we show that the aggregate variables $\{Y^*, K^*, K_E^*, C^*, N^*\}$ and post-tax prices $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ constitute an equilibrium according to Proposition 2 if and only if the aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_N^*\}$ constitute an equilibrium according to Proposition 11.

B.4.1 Alternative characterization of the steady state equilibrium

Proposition 11. *There exists a steady state $\mathscr{S}$ which is consistent with the particular values of aggregate variables aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, prices $\{r_E^*, r_F^*, R_F^*, w^*, \pi_F^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_N^*\}$ and in which no entrepreneurs hide capital or intermediate goods, if and only if the following conditions hold*

1. *The government's budget constraint is balanced every period*

$$\bar{G} = \tau_N^* w^* N^* + \tau_K^* \left(Y^* - w^* N^* - \delta K^*\right) + \tau_W^* K^*. \tag{B.42}$$

2. *Aggregate consumption satisfies*

$$C^* = (\rho + \gamma)\left(K^* + N^* F^{N*} + (1 - N^*) F^*\right). \tag{B.43}$$

We derive aggregate steady state consumption (B.43) by integrating over the consumption policy function of workers and entrepreneurs

$$C^* = (\rho + \gamma) \left(\int_{i \in \mathcal{N}^*} P_i^{N*} di + \int_{i > N^*} P_i^* di \right),$$
$$C^* = (\rho + \gamma) \left(\int_{i \in \mathcal{N}^*} (a_i^{N*} + F^{N*}) di + \int_{i > N^*} (a_i^* + F^*) di \right),$$
$$C^* = (\rho + \gamma)(K^* + N^* F^{N*} + (1 - N^*) F^*),$$

where F^{N*} and F^* are obtained by evaluating $P_{i,t}$ and $P_{i,t}^N$ at steady state prices and steady state tax rates and taking the limit as $\Delta \to 0$, and using $R_{F,t} = 1 + (1 - \tau_K, t)(r_{F,t} - \delta) - \tau_{W,t}$ to simplify

$$F^{N*} = \frac{w^*(1 - \tau_N^*)}{\gamma + \tilde{R}_F^*} \tag{B.44}$$

$$F^* = \frac{(r_E - r_F)(1 - \tau_K^*)\underline{k}_E}{\gamma + \tilde{R}_F^*} \tag{B.45}$$

3. *Aggregate risky capital respects the solution to the entrepreneur's optimization problem, as derived in equation (15):*

$$K_E = \underline{k}_E (1 - N^*) + \mathbb{P} \int_0^1 \hat{k}_E(\theta) \mu^*(\theta) d\theta, \tag{B.46}$$

where $\hat{k}_E(\theta)$ is given by the continuous time policy function (14).

4. *The no-arbitrage condition: entrepreneurs are indifferent between investing in the production of the risk-free intermediate good and lending capital to the bank*

$$\tilde{R}_F^* = (1 - \tau_K^*)(r_F^* - \delta) - \tau_W^*. \tag{B.47}$$

5. *The stationary distribution of entrepreneurial wealth respects transition probabilites and entrepreneur's policy functions*

   (a) *Total entrepreneurial wealth, $\mathbb{P}^*$, is given by*

$$\mathbb{P}^* = \frac{\gamma(1 - N^*)F^*}{\gamma + \rho + (R_F^* - 1) - \int_0^1 (r_E^* - r_F^*)(1 - \tau_K^*)\hat{k}_E(\theta)\mu^*(\theta)\, d\theta} \tag{B.48}$$

(b) *The fraction of wealth owned by entrepreneurs of type* θ, $\mu^*(\theta)$, *is given by*

$$\mu^*(\theta) = \frac{\lambda_\theta h_\theta(\theta) + \gamma h_\theta(\theta) \frac{(1-N^*)F^*}{\mathbb{P}^*}}{\gamma + \rho + \lambda_\theta - \left(\tilde{R}_F^* + \hat{k}_E(\theta)\left(r_E^* - r_F^*\right)\left(1 - \tau_K^*\right)\right)} \tag{B.49}$$

The details of the derivation of the above characterizations of steady state wealth can be found in Appendix B.4.3 below.

6. *Optimality in the production of the final good and market clearning*

$$r_E^* = f_1\left(K_E^*, K^* - K_E^*, N^*\right), \tag{B.50}$$
$$r_F^* = f_2\left(K_E^*, K^* - K_E^*, N^*\right), \tag{B.51}$$
$$w^* = f_3\left(K_E^*, K^* - K_E^*, N^*\right). \tag{B.52}$$

7. *The aggregate fraction of workers* N^* *is determined by the fraction that choose this occupation at birth*

$$N^* = H_z\Bigg(\log((1-\tau_N^*)w^*) - \log((1-\tau_K^*)(r_E^* - r_F^*)\underline{k}_E) - \int_\theta \frac{1}{\rho+\gamma}\left[\frac{(1-\tau_K^*)(r_E^* - r_F^*)\hat{k}(\theta)}{2}\right] dH_\theta(\theta)\Bigg) \tag{B.53}$$

To derive this, we note that the occupational choice implies that in a steady state

$$N^* = H_z((\rho+\gamma)(V^N(F^N, X^*) - \mathbb{E}_\theta V(F, \theta, X^*)). \tag{B.54}$$

In Appendix B.4.4 and B.4.5 we show that the steady state value functions are

$$(\gamma+\rho)V^N\left(\mathbb{P}^{N*}, X^*\right) = \log((\gamma+\rho)\mathbb{P}^N) + \frac{(\tilde{r}_F - \tilde{p}) - \rho}{\gamma+\rho}$$

$$(\gamma+\rho)V(\mathbb{P}, \theta, X^*) = \log\left((\gamma+\rho)\mathbb{P}\right) + \frac{\tilde{R}_F - \rho}{\gamma+\rho} + \frac{1}{\gamma+\rho+\lambda_\theta}\left(\frac{(r_E - r_F)(1-\tau_K)\hat{k}_E(\theta)}{2}\right) + \frac{\lambda_\theta(\gamma+\rho)^{-1}}{\gamma+\rho+\lambda_\theta}\int_0^1 \left(\frac{(r_E - r_F)(1-\tau_K)\hat{k}_E(\theta)}{2}\right) dH_\theta(\theta).$$

Substituting these into (B.54) *yields* (B.53).

8. *The final goods market clears*

$$C^* + \delta K^* + \overline{G} = Y^* = f\left(K_E^* K^* - K_E^* N^*\right). \tag{B.55}$$

9. *The following inequality conditions are satisfied*

$$(r_E^* - r_F^*)(1-\tau_K^*) > 0, \;\; \underline{k}_E(1-N^*) < K_E^* < K^*, \tag{B.56}$$

$$\tilde{R}_F^* + \gamma > 0, \;\; \lambda_\theta + \rho + \gamma - \tilde{R}_F^* > (r_E^* - r_F^*)(1-\tau_K^*)\hat{k}_E(1). \tag{B.57}$$

These follow from entrepreneurial optimization. In particular, equation (B.46) implies that $K_E^* > \underline{k}_E(1-N^*)$ *must hold. In the main text we argued that* $(r_E^* - r_F^*)(1-\tau_K^*) > 0$ *and* $K_E^* < K^*$ *must hold in equilibrium because of the Inada conditions.* $\tilde{R}_F + \gamma > 0$ *must hold, since otherwise (B.45) is inconsistent with* $F^* > 0$*, implying newborn entrepreneurs cannot have positive consumption. The last inequality holds because the denominator of (B.49) must be positive at* $\theta = 1$*, or* $\mu^*(1)$ *would not be positive.*

10. *Lastly,* $\hat{k}_E(\theta)$ *is given by*

$$\hat{k}_E(\theta) = \frac{(r_E^* - r_F^*)(1-\tau_K^*)\theta}{\phi^2(1-\underline{\epsilon})^2\varphi^2}. \tag{B.58}$$

This follows because the inequalities(B.57) and the parameter restrictions imply that

$$\varphi^2 > \lambda_\theta + \rho + 2\gamma > \frac{(r_E^* - r_F^*)(1-\tau_K^*)}{\phi(1-\underline{\epsilon})} \times \min\left[\frac{(r_E^* - r_F^*)(1-\tau_K^*)}{\phi(1-\underline{\epsilon})\varphi^2};1\right],$$

where we used that $\hat{k}_E(1)$ *is given by (14). Dividing by* φ^2*, it follows that*

$$\frac{(r_E^* - r_F^*)(1-\tau_K^*)}{\phi(1-\underline{\epsilon})\varphi^2} < 1.$$

That $\hat{k}_E(\theta)$ *satisfies (B.58) follows from the entrepreneur's problem in (14).*

Note that the asset market clearing condition holds by Walras' law.

B.4.2 Showing equivalence between Proposition 11 and Proposition 2

Proposition 12. *The aggregate variables* $\{K^*, K_E^*, C^*, N^*, Y^*\}$*, functions* $\hat{k}_E(\theta)$*,*$\mu(\theta)$ *and post-tax prices* $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*\right\}$ *constitute a steady state according to Proposition 2 if and only if the aggregate variables* $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$*, , functions* $\hat{k}_E(\theta)$*,*$\mu(\theta)$ *and prices* $\{r_E^*, r_F^*, R_F^*, w^*\}$ *and taxes* $\{\tau_W^*, \tau_K^*, \tau_N^*\}$ *constitute a steady state according to Proposition 11.*

Proof. $\Leftarrow$: Suppose that aggregate variables $\left\{K^*, K_E^*, C^*, N^*, F^*, F^{N*}, \mathbb{P}^*\right\}$, functions $\hat{k}_E(\theta)$,$\mu(\theta)$, prices $\{r_E^*, r_F^*, R_F^*, w^*\}$ and taxes $\{\tau_W^*, \tau_K^*, \tau_N^*\}$ satisfy the equilibrium conditions in Proposition 11: (B.42) - (B.55), and the inequalities (B.56). Define post-tax prices: $\tilde{w} = w(1-\tau_N)$ and $\tilde{r}_X = (r_E - r_F)(1-\tau_K)$.

Replace any occurrence of pre-tax prices in the equilibrium conditions with their post-tax price counterpart, noting that we may replace $(r_E^* - r_F^*)$ with $\tilde{r}_X^*$. The new set of inequality conditions follows directly. By substituting (B.44),(B.45) and (B.48) into the other equilibrium equations to eliminate the variables $\left\{F^*, F^{N*}, \mathbb{P}^*\right\}$, and rearranging, we may arrive at the set of equilibrium conditions of Proposition 2.

⇒: Suppose that aggregate variables $\{K^*, K_E^*, C^*, N^*, Y^*\}$, functions $\hat{k}_E(\theta)$,$\mu(\theta)$ and post-tax prices $\left\{\tilde{r}_X^*, \tilde{R}_F^*, \tilde{w}^*, \tilde{\pi}_F^*\right\}$ satisfy the equilibrium conditions of Proposition 2.

Define r_E^*, r_F^* and w^* according to (B.50)-(B.52); define τ_K^* and τ_N^* so that $\tilde{w} = w(1-\tau_N)$ and $\tilde{r}_X = (r_E - r_F)(1-\tau_K)$; define τ_W according to (B.47); define $\left\{F^*, F^{N*}, \mathbb{P}^*\right\}$ according to (B.44),(B.45) and (B.48). Then, it follows by inspection that the resulting variables satisfy the equilibrium conditions of Proposition 11. □

B.4.3 Derivation of $\mathbb{P}^*$ and $\mu^*(\theta)$ (equations (B.48) andb (B.49))

Derivation of $\mathbb{P}^*$. Consider the model with period length Δ. In a steady state, each period a fraction $\gamma\Delta$ of entrepreneurs die and a fraction $(1 - N^*)$ of the $\gamma\Delta$ newborn agents choose to become entrepreneurs, each with initial wealth F. The law of motion for $\mathbb{P}$ is then

$$\mathbb{P}' = (1 - \gamma\Delta) \int_{i \in \mathcal{N}} P_i di + \gamma\Delta(1 - N^*)F.$$

This can be written as

$$\mathbb{P}' = (1 - \gamma\Delta)(1 - \rho\Delta) \left[(r_E^* - r_F^*)\Delta(1 - \tau_K^*) \int_0^1 \hat{k}_E(\theta)\mu^*(\theta)\mathrm{d}\theta + (1 + \tilde{R}_F^*\Delta)\right] \mathbb{P} + \gamma\Delta(1 - N)F,$$

where we evaluate the integral $\int_{i>N} P_i di$ by combining the expression ω (equation (B.17)) and the entrepreneur's policy function for $\mathbb{P}^*$ (equation (B.19)) and the entrepreneur's policy function for k_E (equation (B.16)). Using that in steady state $\mathbb{P}' = \mathbb{P} = \mathbb{P}^*$ and taking the limit as $\Delta \to 0$, we obtain equation (B.48), as required.

Derivation of $\mu^*(\theta)$. Recall that we define $\mu^*(\theta)$ as the fraction of total entrepreneur wealth, $\mathbb{P}^*$, held by entrepreneurs of type θ in the steady state

$$\mu^*(\theta) := \frac{\int_{i \in \mathcal{N}^*, \theta_i = \theta} P_i^* di}{\mathbb{P}^*} \tag{B.59}$$

The law of motion for $\mu'(\theta)$ in an economy of length Δ is given by

$$\begin{aligned} \mu'(\theta)\mathbb{P}' =& \lambda_\theta \Delta h_\theta(\theta)(\mathbb{P}' - \gamma\Delta(1 - N)F) + \gamma\Delta g(\theta)(1 - N)F \\ &+ (1 - \gamma\Delta)(1 - \lambda_\theta\Delta)\mu(\theta)\mathbb{P} \int_\xi \hat{P}'(\theta, \xi)\mathrm{d}H_\xi(\xi) \end{aligned} \tag{B.60}$$

At the end of each period, entrepreneurs die with probability $\gamma\Delta$. Out of the $(1-\gamma\Delta)$ entrepreneurs that survive, they keep the same θ with probability $(1-\lambda_\theta\Delta)$. The left hand side corresponds to the total capital held by entrepreneurs of type θ in the next period. The three right hand terms are derived as follows

1. The first term indicates the capital held in the next period by surviving entrepreneurs who draw a new ability level that period and happen to draw θ. In total, surviving entrepreneurs hold $\mathbb{P}' - \gamma\Delta(1-N)F$ of capital. Since the probability of drawing a new ability level is orthogonal to ability type, the fraction of entrepreneurs drawing theta as their new type is equal to the fraction of wealth these entrepreneurs own.

2. The second term is given by the wealth owned by newborns, $(1-N)F$, scaled by the fraction of newborns who happen to draw θ.

3. Fraction $(1-\gamma\Delta)(1-\lambda_\theta\Delta)$ of entrepreneurs of type θ surive and retain their type in the next period. Due to the linearity of the policy function in P, the cumulative wealth of these entrepreneurs in the next period is equal to the expected wealth of an entrepreneur who owns their cumulative wealth, $(1-\gamma\Delta)(1-\lambda_\theta\Delta)\mu(\theta)\mathbb{P}$. Note that we have defined $\hat{P}'(\theta,\epsilon) := \frac{P'}{P}$.

In a steady state, $\mu' = \mu = \mu^*$ and $\mathbb{P}' = \mathbb{P} = \mathbb{P}^*$ and equation (B.60) simplifies to

$$\mu^*(\theta) = \frac{\lambda_\theta \Delta h_\theta(\theta) + \gamma\Delta(1-N^*)\frac{F^*}{\mathbb{P}^*}h_\theta(\theta)(1-\lambda_\theta\Delta)}{1-(1-\gamma\Delta)(1-\lambda_\theta\Delta)\int_\xi \hat{P}'(\theta,\xi)\mathrm{d}H_\xi(\xi)} \tag{B.61}$$

To find the expression for $\int_\epsilon \hat{P}'(\theta,\epsilon)\mathrm{d}H(\epsilon)$, we substitute the expression for ω given by equation (B.17) into the entrepreneur's policy function (B.19). We then use the fact that k_E is linear in P and that $\mathbb{E}[\epsilon] = 1$ to show that

$$\begin{aligned}
\int_\xi \hat{P}'(\theta,\epsilon)\mathrm{d}H(\epsilon) &= (1-\rho\Delta)\int_\epsilon \omega(\theta,\xi,1)\mathrm{d}H_\xi(\xi) \\
&= (1-\rho\Delta)\int_\xi \Big((\phi(\epsilon-1) + (1-\tau_K)(r_E - r_F)\Delta)\hat{k}_E(\theta) + (1+\tilde{R}_F\Delta)\Big)\mathrm{d}H_\xi(\xi) \\
&= (1-\rho\Delta)\Big((r_E - r_F)(1-\tau_K)\Delta\hat{k}_E(\theta) + (1+\tilde{R}_F\Delta)\Big)
\end{aligned}$$

Taking the limit as $\Delta \to 0$, we obtain equation (B.49), as required.

B.4.4 Derivation of the worker's steady state value function

In this section, we show that that the worker's steady state value function is given by (B.55).

We derive this continuous time Bellman equation by taking the limit of the Bellman equation of the problem for a given period length Δ

$$V^N(P^N, X^*) = \max\left(\log(c^N)\Delta + (1-\rho\Delta)(1-\gamma\Delta)\, V^N\left(P^{N\prime}, X^*\right)\right)$$

$$(\rho+\gamma-\rho\gamma\Delta)V^N(P^N, X^*) = \max\left(\log(c^N) + (1-\rho\Delta)(1-\gamma\Delta)\,\frac{V^N\left(P^{N\prime}, X^*\right) - V^N(P^N, X^*)}{\Delta}\right)$$

Hence, taking the limit as $\Delta \to 0$, and using equation (14) and substituting c^N, we obtain

$$(\rho+\gamma)V^N(P^N, X^*) = \log((\rho+\gamma)P^N) + \underbrace{\lim_{\Delta\to 0}\left(\frac{V^N\left(P^{N\prime *}, X^*\right) - V^N(P^N, X^*)}{\Delta}\right)}_{\text{A}} \tag{B.62}$$

As the next step, we compute the term A. As in Appendix A.1, we can write the worker's steady state value function in an economy with period length Δ as

$$V^N(P^N, X^*) = Q(\Delta) + \frac{1}{1-(1-\rho\Delta)(1-\gamma\Delta)}\log(P^N)\Delta. \tag{B.63}$$

Using the policy function for P' for the economy of period length Δ, given by equation (B.9), and l'Hôpital's rule we obtain

$$\begin{aligned} A &= \lim_{\Delta\to 0}\left(\frac{1}{1-(1-\rho\Delta)(1-\gamma\Delta)}\log\Big((1-\rho\Delta)(1+(R_F-1)\Delta)\Big)\right) \\ &= \lim_{\Delta\to 0}\left(\frac{1}{\rho+\gamma-2\rho\gamma\Delta}\,\frac{-\rho(1+(R_F-1)\Delta)+(1-\rho\Delta)(R_F-1)}{(1-\rho\Delta)(1+(R_F-1)\Delta)}\right) = \frac{\tilde{R}_F-\rho}{\rho+\gamma}. \end{aligned}$$

B.4.5 Derivation of the entrepreneur's steady state value function

In this section we solve for the entrepreneur's steady state value function. We evaluate equation (B.4) at $X = X' = X^*$ and substitute in the policy functions for c and P'

$$\begin{aligned} V(P,\theta,X^*) \;&=\; \int_\xi \Bigg(\log(c(P,\theta,\epsilon,X^*)))\Delta + (1-\rho\Delta)(1-\gamma\Delta) \\ &\times\; E\left[V(P'(P,\theta,\epsilon,X^*)),\theta',X^*)\middle|\theta\right]\Bigg)\mathrm{d}H_\xi(\xi) \end{aligned} \tag{B.64}$$

Using the process for θ, we can write out the conditional expectation as

$$E\left[V(P',\theta',X^*)\middle|\theta\right] = [1-\lambda_\theta\Delta]V(P',\theta,X^*) + \lambda_\theta\Delta\mathbb{E}_\theta[V(P',\theta,X^*)]$$

As in Lemma 1, we write the entrepreneur's value function when the period length is Δ as

$$V(P,\theta,X)=\bar{V}(\theta,X,\Delta)+\frac{1}{1-(1-\rho\Delta)(1-\gamma\Delta)}\log(P)\Delta. \tag{B.65}$$

Hence, as $\Delta\to 0$, we obtain the continuous time entrepreneur value function

$$V(P,\theta,X^*)=\lim_{\Delta\to 0}\bar{V}(\theta,X^*,\Delta)+\frac{1}{\rho+\gamma}\log(P). \tag{B.66}$$

It remains to derive $\bar{V}(\theta,X,\Delta)$ and take the limit as $\Delta\to 0$. Since we have already derived the solutions for c, P' and k_E, deriving $\bar{V}(\theta,X,\Delta)$ is simply a matter of substituting these expressions into the value function and rearranging. To simplify the algebra, define $\alpha_\Delta:=(1-\rho\Delta)(1-\gamma\Delta)$ to obtain

$$\omega=\underline{\omega}+\theta\epsilon k_E \underset{(B.17),(B.16)}{=} \underbrace{\Big((\phi(\epsilon-1)+(1-\tau_K)(r_E-r_F)\Delta)\hat{k}_E(\theta)+(1+\tilde{R}_F\Delta)\Big)}_{=:D(\theta,\epsilon,\Delta)}P \tag{B.67}$$

$$c\underset{(B.10)}{=}(\rho+\gamma+\rho\gamma\Delta)\omega=(\rho+\gamma+\rho\gamma\Delta)D(\theta,\epsilon,\Delta)P, \tag{B.68}$$

$$P'\underset{(B.11)}{=}(1-\rho\Delta)\omega=(1-\rho\Delta)D(\theta,\epsilon,\Delta)P \tag{B.69}$$

Combining the closed form solution for the value function (B.65) with the Bellman equation (B.64), and substituting in for c and P' with the expressions above, we obtain

$$\begin{aligned}\bar{V}(\theta,X^*,\Delta)+\frac{1}{1-\alpha_\Delta}\log(P)\Delta=\int_\xi\bigg(&\log\left((\rho+\gamma+\rho\gamma\Delta)D(\theta,\epsilon,\Delta)P\right)\Delta+\alpha_\Delta\\ \times\bigg\{[1-\lambda_\theta\Delta]\Big(&\bar{V}(\theta,X^*,\Delta)+\frac{1}{1-\alpha_\Delta}\log\left((1-\rho\Delta)D(\theta,\epsilon,\Delta)P\right)\Delta\Big)\\ +\lambda_\theta\Delta\Big(&\mathbb{E}_\theta[\bar{V}(\theta,X^*,\Delta)]+\frac{1}{1-\alpha_\Delta}\log\left((1-\rho\Delta)D(\theta,\epsilon,\Delta)P\right)\Delta\Big)\bigg\}\bigg)\mathrm{d}H(\epsilon).\end{aligned}$$

This simplifies to the following expression for $\bar{V}(\theta,X^*,\Delta)$

$$\begin{aligned}(1-\alpha_\Delta(1-\lambda_\theta\Delta))\bar{V}(\theta,X^*,\Delta) &= \log(\rho+\gamma+\rho\gamma\Delta)\Delta+\frac{\alpha_\Delta\Delta}{1-\alpha_\Delta}\log(1-\rho\Delta)\\ &+ \alpha_\Delta\lambda_\theta\Delta\mathbb{E}_\theta[\overline{V}(\theta,X^*,\Delta)]+\frac{\Delta}{1-\alpha_\Delta}\int_\xi\log(D(\theta,\epsilon,\Delta))dH_\xi(\xi)\end{aligned}$$

Taking expectations and rearranging, we have

$$\begin{aligned}\mathbb{E}_\theta \bar{V}(\theta, X^*, \Delta) &= \frac{\log(\rho+\gamma+\rho\gamma\Delta)\Delta}{1-\alpha_\Delta} + \frac{\alpha_\Delta \Delta}{(1-\alpha_\Delta)^2}\log(1-\rho\Delta) \\ &+ \frac{\Delta}{(1-\alpha_\Delta)^2}\mathbb{E}_\theta\left[\int_\xi \log(D(\theta,\epsilon,\Delta))dH_\xi(\xi)\right]\end{aligned}$$

Substituting this into the previous expression and dividing both sides by Δ, we obtain

$$\begin{aligned}&\left(\frac{1-\alpha_\Delta}{\Delta} + \alpha_\Delta\lambda_\theta\right)\bar{V}(\theta, X^*, \Delta) \\ = &\left(1+\alpha_\Delta\lambda_\theta\frac{\Delta}{1-\alpha_\Delta}\right)\left[\log(\rho+\gamma+\rho\gamma\Delta) + \alpha_\Delta\left(\frac{\Delta}{1-\alpha_\Delta}\right)\frac{\log(1-\rho\Delta)}{\Delta}\right] \\ + &\alpha_\Delta\lambda_\theta\left(\frac{\Delta}{1-\alpha_\Delta}\right)^2\mathbb{E}_\theta\left[\frac{\int_\xi \log(D(\theta,\epsilon,\Delta))dH_\xi(\xi)}{\Delta}\right] + \left(\frac{\Delta}{1-\alpha_\Delta}\right)\frac{\int_\xi \log(D(\theta,\epsilon,\Delta))dH_\xi(\xi)}{\Delta}\end{aligned} \tag{B.70}$$

It remains to take the limit as $\Delta \to 0$. The following limits obtain

$$\lim_{\Delta\to 0}\log(\rho+\gamma+\rho\gamma\Delta) = \log(\rho+\gamma) \tag{B.71}$$

$$\lim_{\Delta\to 0}\alpha_\Delta = 1 \tag{B.72}$$

$$\lim_{\Delta\to 0}\frac{\Delta}{1-\alpha_\Delta} = \frac{1}{\rho+\gamma} \tag{B.73}$$

$$\lim_{\Delta\to 0}\frac{\log(1-\rho\Delta)}{\Delta} = -\rho \tag{B.74}$$

It is immediate that equation (B.71) holds. Equation (B.72) is immediate given that $\alpha_\Delta = (1-\rho\Delta)(1-\gamma\Delta)$. Then, equations (B.73) and (B.74) follow from L'Hopital's rule.

We prove below that the following limit also obtains

$$\lim_{\Delta\to 0}\left[\frac{\int_\xi \log(D(\theta,\epsilon,\Delta))dH_\xi(\xi)}{\Delta}\right] = (r_E - r_F)(1-\tau_K)\hat{k}(\theta) + \tilde{R}_F - \frac{1}{2}\left(\frac{\phi(1-\underline{\epsilon})\hat{k}(\theta)\varphi}{\theta}\right)^2 \tag{B.75}$$

We now take the limit of (B.70), using (B.71)-(B.75). We obtain

$$\begin{aligned}(\rho+\gamma+\lambda_\theta) \quad \lim_{\Delta\to 0}\bar{V}(\theta, X^*, \Delta) &= \left(1+\frac{\lambda_\theta}{\rho+\gamma}\right)\left[\log(\rho+\gamma) + \left(\frac{1}{\rho+\gamma}\right)(-\rho)\right] \\ &+ \lambda_\theta\left(\frac{1}{\rho+\gamma}\right)^2\mathbb{E}_\theta\left[(r_E - r_F)(1-\tau_K)\hat{k}(\theta) + \tilde{R}_F - \frac{1}{2}\left(\phi(1-\underline{\epsilon})\hat{k}(\theta)\varphi\right)^2\right] \\ &+ \left(\frac{1}{\rho+\gamma}\right)\left[(r_E - r_F)(1-\tau_K)\hat{k}(\theta) + \tilde{R}_F - \frac{1}{2}\left(\phi(1-\underline{\epsilon})\hat{k}(\theta)\varphi\right)^2\right].\end{aligned}$$

Combining this with (B.66), using (B.58) and rearranging, we obtain (B.55). It remains only to show that the limit (B.75) obtains. Using the definition of $D(\theta, \epsilon, \Delta)$, we have

$$\begin{aligned}
&\lim_{\Delta\to 0}\frac{1}{\Delta}\int_{\xi}\log\left(D(\theta,\epsilon,\Delta)\right)\mathrm{d}H_{\xi}(\xi)\\
&=\lim_{\Delta\to 0}\frac{1}{\Delta}\int_{\xi}\log\left(\left[\phi(\epsilon-1)+(r_E-r_F)\Delta(1-\tau_K)\right]\hat{k}_E(\theta)+(1+\tilde{R}_F\Delta)\right)\mathrm{d}H_{\xi}(\xi)\\
&=\lim_{\Delta\to 0}\frac{1}{\Delta}\int_{\xi}\log\left(\left[\phi(1-\underline{\epsilon})\left(\exp\left(\frac{\varphi\sqrt{\Delta}\xi}{\sqrt{\theta}}-\frac{\varphi^2\Delta}{2\theta}\right)-1\right)+(r_E-r_F)\Delta(1-\tau_K)\right]\hat{k}_E(\theta)\right.\\
&\left.+(1+\tilde{R}_F)\right)\mathrm{d}H_{\xi}(\xi)
\end{aligned}$$

We now take the Taylor expansion of the integrand in units of $\sqrt{\Delta}$ around the point $\sqrt{\Delta}=0$. To do this let $a=\phi(1-\underline{\epsilon})\hat{k}$, $b=\frac{\varphi\xi}{\sqrt{\theta}}$, $c=\frac{\varphi^2}{2\theta}$, $d=(r_E-r_F)(1-\tau_K)\hat{k}+\tilde{R}_F$. Then, we can write the integrand as $f(\sqrt{\Delta})$, where $f(\sqrt{\Delta})=\log(a(\exp(b\Delta-c\Delta)-1)+d\Delta+1)$ and so

$$\begin{aligned}
f'(\sqrt{\Delta}) &= \frac{a(b-2c\sqrt{\Delta})\exp(b\sqrt{\Delta}-c\Delta)+2d\sqrt{\Delta}}{a(\exp(b\sqrt{\Delta}-c\Delta)-1)+d\Delta+1}\\
f''(\sqrt{\Delta}) &= (a(\exp(b\sqrt{\Delta}-c\Delta)-1)+d\Delta+1)^{-2}\\
&\times \Big\{(a(\exp(b\sqrt{\Delta}-c\Delta)-1)+d\Delta+1)\exp(b\sqrt{\Delta}-c\Delta)(-2ac+a(b-2c\sqrt{\Delta})^2+2d)\\
&- (a(b-2c\sqrt{\Delta})\exp(b\sqrt{\Delta}-c\Delta)+2d\sqrt{\Delta})^2\Big\}
\end{aligned}$$

Then, the Taylor expansion of the integrand above around $\sqrt{\Delta}=0$ is as follows

$$\frac{\varphi\xi\phi(1-\underline{\epsilon})\hat{k}\sqrt{\Delta}}{\sqrt{\theta}}+\frac{\frac{\Delta}{2}(\xi^2-1)\phi(1-\underline{\epsilon})\hat{k}\varphi^2}{\theta}+\Delta(r_E-r_F)(1-\tau_K)\hat{k}+\Delta\tilde{R}_F-\frac{\Delta}{2}\left(\frac{\phi(1-\underline{\epsilon})\hat{k}\varphi\xi}{\sqrt{\theta}}\right)^2.$$

Integrating this over ξ and using that $E[\xi]=0$ and $E[\xi^2]=1$, we therefore conclude that

$$\begin{aligned}
&\lim_{\Delta\to 0}\int_{\xi}\frac{1}{\Delta}\log\left(D(\theta,\epsilon,\Delta)\right)\mathrm{d}H_{\xi}(\xi)\\
&\quad=\lim_{\Delta\to 0}\frac{1}{\Delta}\left(\Delta(r_E-r_F)(1-\tau_K)\hat{k}+\Delta\tilde{R}_F-\frac{\Delta}{2\theta}\left(\phi(1-\underline{\epsilon})\hat{k}\varphi\right)^2+\mathcal{R}\right)\\
&\quad=(r_E-r_F)(1-\tau_K)\hat{k}+\tilde{R}_F-\frac{1}{2\theta}\left(\phi(1-\underline{\epsilon})\hat{k}\varphi\right)^2,
\end{aligned}$$

where $\mathcal{R}$ are some terms of higher order in Δ. This is the same as equation (B.75). □.

B.5 Characterizing Partial Equilibrium Elasticities of Tax Changes

To obtain Propositions 3-5, we precisely 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.

We formally define $\frac{\partial Y}{\partial \tau_K}$ using Proposition 2, equation (B.47) and the definition of $\tilde{r}_X$. First, differentiate the latter two with respect to τ_K, holding constant pre-tax prices to obtain

$$\frac{\partial \tilde{r}_X}{\partial \tau_K} = (r_E - r_F) \quad \text{and} \quad \frac{\partial \tilde{R}_F}{\partial \tau_K} = r_F - \delta.$$

Then, we make use of Proposition 2 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 $\tilde{R}_F$.[25] Therefore, we cantreat K, Y, K_E and C as functions of $\tilde{r}_X$ and $\tilde{R}_F$, and use the equations of Proposition 2 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 $\frac{\partial \tilde{R}_F}{\partial \tau_K}$ found above, 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 steady state variable X, and for $j \in \{K; W\}$

$$e^X_{\tau_K} = \frac{1-\tau_K}{X}\frac{\partial X}{\partial \tau_K} = -\frac{\tilde{r}_X}{X} \cdot \frac{\partial X}{\partial \tilde{r}_X} - \frac{(1-\tau_K)(r_F - \delta)}{X}\frac{\partial X}{\partial \tilde{R}_F} \tag{B.76}$$

$$e^X_{\tau_W} = \frac{1}{X}\frac{\partial X}{\partial \tau_W} = -\frac{1}{X}\frac{\partial X}{\partial \tilde{R}_F}, \tag{B.77}$$

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 Proof of Proposition 3

Differentiate $Y = f(K_E, K, N)$ with respect to $\tilde{r}_X$ and $\tilde{R}_F$, using that the derivatives of f are given by (B.50)-(B.52). Thus, for $x \in \{\tilde{r}_X; \tilde{R}_F\}$

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

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

Substitute $\frac{\partial Y}{\partial \tilde{r}_X}$ and $\frac{\partial Y}{\partial \tilde{R}_F}$ into equations (B.76) and (B.77). The result follows immediately.

B.5.2 Proof of Proposition 4

Differentiate (19) with respect to $\tilde{r}_X$ and $\tilde{R}_F$. Substitute into equations (B.76) and (B.77) and rearrange. The result follows.

B.5.3 Lemma 3 Proof

Given the definitions of $e^{\hat{k}}_{\tau_j}$, for $j \in \{K; W\}$, in Proposition 4, it is sufficient to show that

$$\lim_{\lambda_\theta \to \infty} (1-\tau_K) \frac{\partial}{\partial \tau_K} \log \left(\int_\theta \mu(\theta) \hat{k}_E(\theta) \mathrm{d}\theta \right) = -1 \quad \text{and} \quad \lim_{\lambda_\theta \to \infty} \frac{\partial}{\partial \tau_W} \log \left(\int_\theta \mu(\theta) \hat{k}_E(\theta) \mathrm{d}\theta \right) = 0.$$

Using (21), 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 (20), 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 (21), 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.5.4 Proof of Proposition 5

Let $x \in \{\tilde{r}_X; \tilde{R}_F\}$. Combining equations (16) and (17) 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 3 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.76) and (B.77) rearranging, we obtain the desired result.

B.6 Effects of Tax Changes on Worker Lifetime Utility

We focus first on the discrete time case, before moving to the continuous time case discussed in Lemma 4 . 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.[26] 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

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

[26] 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 we have

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

Using that the worker consumes $c_0^N = [1-(1-\rho)(1-\gamma)](1+\tilde{R}_F)F^N$, 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, we obtain

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

where

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

Simplifying and taking the limit as $\Delta \to 0$, we obtain the result of Lemma 4.

B.7 Proof of Proposition 9

Repeating the same steps as used in Appendix B.6, we obtain that the change in a newborn worker's lifetime utility from a change in taxes in the continuous time case is given by

$$\mathrm{d}V^N(0, X) = u'(c_0^N)\left(\frac{F^N}{\tilde{w}}\right)\left[\mathrm{d}\tilde{w} + \mathcal{A}^N\mathrm{d}\tilde{R}_F\right], \tag{B.78}$$

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 $u'(c_0^N)$ is no longer the same as $\frac{1}{c_0^N}$.

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

$$\mathrm{d}N = H_z'(z^*)[\mathrm{d}V^N(0,X) - \mathrm{d}\mathbb{E}_\theta V_E(0,\theta,X)] \tag{B.79}$$

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

$$e_{\tilde{w}}^N = \frac{\tilde{w}}{N}\frac{\partial N}{\partial \tilde{w}} = \frac{H_z'(z^*)\tilde{w}}{N}\frac{\partial V^N}{\partial \tilde{w}}$$

Combining this with equation (B.78) and rearranging, we obtain

$$\mathrm{d}V^N(0,X) = \frac{Ne_{\tilde{w}}^N}{H_z'(z^*)\tilde{w}}\left[\mathrm{d}\tilde{w} + \mathcal{A}^N \mathrm{d}\tilde{R}_F\right] \tag{B.80}$$

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.79) and (B.80), we obtain

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

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

$$\mathrm{d}\mathcal{W} = \frac{e_{\tilde{w}}^N}{H_z'(z^*)\tilde{w}}\left[-w\mathrm{d}\tau_N N - B_{\tau_j}^N N\mathrm{d}\tau_j - \frac{(1-N)w(1-\tau_N)\mathrm{d}N}{e_{\tilde{w}}^N}\right].$$

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, as done in the derivation of Proposition 7 in the main text. 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 8, 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 8.

Rearranging this to solve for optimal taxes and writing in matrix form, we arrive at the optimal tax formula in Proposition 8. □

B.8 Values of Terms in the Optimal Tax Formula

The remaining terms used in the Optimal Tax formula in Proposition 8 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.

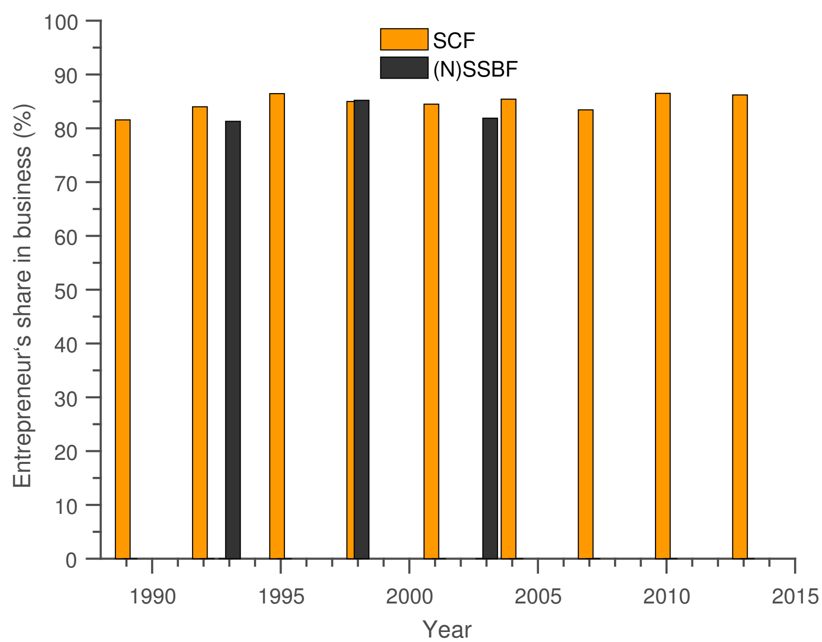
The Survey of Consumer Finances is a a triennial cross-sectional survey of U.S. families which provides information on individual household portfolio composition, including investment in private firms. While the SCF was initially administered in 1983, it was not until

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.[27]

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 shareholder. 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.

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