Wealth Inequality, Uninsurable Entrepreneurial Risk and Firms Markup

Samuel Brien
Queen's University

November 28, 2021

Abstract

This paper examines the effect of wealth concentration on firms' market power when firm entry is driven by entrepreneurs facing uninsurable idiosyncratic risks. Under greater wealth concentration, households in the lower end of the wealth distribution are more risk averse and less willing (or able) to bear the risk of entrepreneurial activities. This has implications for firm entry, competitiveness, and market power.

I calibrate a Schumpeterian model of endogenous growth with heterogeneous risk averse entrepreneurs competing to catch up with firms. This model is unique in that both household wealth distribution and a measure of firm markup are endogenously determined on a balanced growth path. I find that a spread in the wealth distribution decreases entrepreneurial firm creation, resulting in greater aggregate firm market power. This result is supported by time series evidence obtained from the estimation of a structural panel VAR with OECD data from eight countries.

Keywords: Wealth inequality, market power, growth, Schumpeterian, endogenous growth, entrepreneur.

JEL Codes: E220, E250, L120, O310, O330, O340.

Acknowledgments: I am grateful to Brant Abbott, Morten Ø. Nielsen for their supervision, insightful comments and suggestions. I want to thank Allen Head, James MacKinnon and other seminar participants at Queen's University Economics Department for their useful comments, as well as participants at the 2021 annual conference of the *Société canadienne de science économique* (SCSE). I gratefully acknowledge the financial support given by the Social Sciences and Humanities Research Council of Canada (SSHRC).

1 Introduction

Over the last few decades , U.S. business dynamism has been on the decline, as job creation, labour share, capital investment and output growth have followed a decreasing trend (Akcigit and Ates, 2019; Eggertsson et al., 2018, 2019). Concurrently, wealth and income inequalies have also risen alongside firms' price-cost markups and other indicators of market power like firm concentration indices and the Tobin's Q (Philippon, 2019; Gutierrez and Philippon, 2017) [1].

Recent studies point to firms' market power as a culprit for the current sluggishness of Western economies. The apparent increase in market power is usually attributed to a decrease in the knowledge diffusion rate (Akcigit and Ates, 2019), lax antitrust policies (Philippon, 2019), incomplete passthrough of technological efficiency gains (De Loecker et al., 2021) or the reallocation of resources towards more profitable firms (Autor et al., 2020; De Loecker et al., 2020). An unexplored line of inquiry is the effect of wealth inequality on market power through reduced entrepreneurial activity and firm entry. The present work fills that gap.

I propose a Schumpeterian model of endogenous growth where firm entry is the consequence of entrepreneurial activity. Risk averse households choose their level of engagement in a risky entrepreneurial activity and the uninsurable nature of the entrepreneurial risk is such that relatively poorer households are not as willing to borrow as they would if they could be insured, while wealthier households face a decreasing marginal return on their entrepreneurial effort. The aggregate entrepreneurial activity drives firms entry, competition and market power.

I calibrate the model with U.S. data. The framework generates both an endogenous distribution of household wealth and an endogenous measure of market power. It is thus well-suited to examine the relationship between the two measures, without making a priori assumption on the direction of the causal links, if any.

This study presents two main findings. First, wealth concentration decreases the rate of firm entry and increases market power. This finding is generated by the theoretical model irrespective of the driver of inequality, and it is supported by time series evidence obtained from a structural panel vector autoregression (PVAR) estimation of yearly data from eight OECD countries. Using De Loecker et al. (2020) markup estimate as an indicator of market power, we find that the response of markup to a shock of inequality propagating via the channel of firm entry is positive and significant at a 15% level. Results are robust to a selection of inequality indicators and the variations in inequality explains about 5% of the

[1] Although there is currently a heated debate as to the best methodology for markup estimation and their interpretation, there is little disagreement to the effect that markups are higher now than they were in the early 1980s. I discuss markups estimates in Section 2.1.

variance of firm markups. The empirical exercise is described in detail in Section 2.

The second result is in line with the more conventional way of thinking about market power and inequality: market power exacerbates wealth inequality as suggested by Comanor and Smiley (1975), Ennis et al. (2019), Furman and Orszag (2015), Stiglitz (2017) and Brun and González (2017). Both the theoretical model and the empirical exercise find that inequality responds positively to a variation in market power and markup, respectively, with variations of markup explaining about 20% of the inequality indicator's variance. This second result is not new, however; thus, the focus of this paper will be on the first result which is new to the literature.

Foellmi and Zweimüller (2004, 2006) also propose a model where household wealth spread is positively associated with firms' market power. In their framework, the mechanism propagates through a differentiated consumer demand channel. I propose a channel of household entrepreneurial activity. To my knowledge, this is the first attempt to model the effect of household wealth on firms' market power through an occupational choice mechanism. In a closely related work, Banerjee and Newman (1994) propose a model of imperfect credit market in which household occupational decisions are dependent upon wealth. Their model implies that only wealthy agents become entrepreneurs and contribute to growth. The model herein proposed differs from Banerjee and Newman (1994) in that entrepreneurs face uninsurable risk instead of borrowing impediments. Also, entrepreneurs do not contribute to growth directly, but through increased firm competition induced by entrepreneurial entry.

The main contribution of this paper lies in the proposed general equilibrium framework featuring the incorporation of heterogeneous risk averse entrepreneurs to a step-by-step innovation framework of endogenous growth (Aghion et al., 2001) proposed by (Akcigit and Ates, 2019). Solving the entrepreneurs problem is made possible with the normalization presented in Appendix B.1.3 and an adaptation of the finite-difference algorithm offered in Achdou et al. (2020). García-Peñalosa and Wen (2008) also incorporate occupational choice to a similar Schumpeterian model of endogenous growth; however, they solve a partial equilibrium taking the interest rate as given.

Besides the new theoretical framework, this papers offers two methodological contributions to the time series literature in the empirical exercise: an application of a bootstrap-based bias correction approach to the PVAR estimation and a new counterfactual decomposition method for PVAR impulse response functions (IRF). A particular challenge faced in the PVAR estimation came from the short dynamic panel and the resulting Hurwicz-type bias (Hurwicz, 1950). Because of the presence of a trend in the model and for other reasons detailed in Section 2.2, existing PVAR unbiased estimators are not well-suited to this problem. I perform Monte-Carlo simulations to show that despite the large number of parameters to estimate, MacKinnon and Smith (1998) bootstrap-based bias-correction approach, when applied to the maximum likelihood (ML) estimator, performs remarkably well in this dynamic panel setting.

The time series exercise also features a new IRF channel decomposition method, estimating the response of firm markup to an inequality shock that propagates via the firm entry channel. The method is equivalent to that of Sims and Zha (2006) and Bernanke et al. (1997) channel shutoff method, but has the advantage of being obtained with a simple calculation using quantities that are usually already computed when reporting IRFs.

The entrepreneurial activity channel at the core of the theory presented in this paper provides important insight for policy. I examine the effect of five fiscal policies on wealth inequality, firm entry, market power and growth. The policies considered are three taxes on capital gains, corporate profits and wealth, and two subsidies on entrepreneurial and firm R&D expenditures. Consistent with work by Cozzi (2018) and Conesa et al. (2009), I find that under the calibrated benchmark, a slight wealth tax is growth enhancing, due to the uninsurable nature of entrepreneurs risk. I decompose the effect of each policy into an incentive effect that modifies agents' first-order conditions conditional on their wealth and a redistributive effect that stems from the change in the shape of the wealth distribution. I find that the incentive effect dominates everywhere except in the case where a small wealth tax is applied.

Both the wealth tax and the entrepreneurial subsidy decrease wealth inequality and market power and stimulate firm entry and growth. It is worth noting that the firm R&D subsidy also stimulates growth, but at the expense of greater firm market power and wealth concentration. Macroeconomics indicators elasticities to the five policies are discussed in Section 5.3.

The remainder of the paper is structured as follows. The project is motivated by time series evidence offered in Section 2. The framework is then presented in Section 3, followed by a calibration exercise and a discussion of the main takeaway from the model in Sections 4 and 5 before concluding.

2 Time Series Evidence

In this section, I examine the comovements between indicators of household inequality, entrepreneurial activity and firm price-cost markups. I find that a variation of inequality is significantly associated with a future rise in the markup indicator and that a significant part of the effect propagates through the entrepreneurial activity indicator. The effect is significant, but is not very large, explaining at most 5% of the variations in markup. The evidence is also consistent with the converse relation (Furman and Orszag, 2015; Stiglitz, 2017) as firm markups explain around 20% of the variations of the inequality indicator.

In the following section, I discuss the data, the challenges associated with the short panel and the bias-correction method that I use. Presentation of estimation results follows.

2.1 Data

A structural panel VAR using yearly data is estimated. Included in the unbalanced panel are eight countries chosen for the availability of uninterrupted series on entrepreneurial activity: Canada, Finland, France, Italy, Netherlands, Norway, the U.S. and the U.K. The four variables examined are a short-term interest rate, the rate of entry of new firms, the share of income of the top 1% of households and De Loecker et al. (2020) markup estimates. Figure 1 displays the data and the different periods covered.

Short-term interest rates based on three-month money market rates, and firm entry rates are obtained from the OECD database. The rate of entry of new firms is defined as the ratio of new employer firms over the total number of such firms. The data is obtained from local business registries, and the exact definition may differ slightly between countries[2]. Unfortunately, the firm entry series are rather short and a usable series starting before the early 2000s is only available for eight OECD countries. I discuss the challenges associated with this short and narrow panel in the methodology section below.

While it would be ideal for wealth data for the OECD countries above to be available, this is unfortunately not the case. I follow Aghion et al. (1999) and use the share of income earned by households above the 99th percentile as a proxy for household wealth inequality.

As a proxy for market power, I use the firm markup estimates provided by De Loecker et al. (2020). It is worth keeping in mind that the apparent increase in markups is not necessarily indicative of an increase in firms market power. As pointed out by the authors, the share of administrative and overhead costs have also increased in the last decades, and it is not clear to what extent the increase in markup is a reflection of an increase in the share of fixed costs.

[2]For instance, the Canadian firm entry rate only includes firms with less than 250 employees.

I use De Loecker et al. (2020)'s markup estimates because they seem to be the best estimates readily available for the eight OECD countries included in the panel. These estimates nonetheless have drawn some criticism. Three main critiques are made. First, Traina (2018) points out that administrative, marketing and overhead firm costs have evolved in the last decades and argues that, consequently, they should be included in the measure of variable costs. This is a matter of definition; as long as one understands what is being measured, the estimate remains informative whether or not administrative costs are included. De Loecker et al. (2020) concede that their markup estimates are not necessarily indicative of market power. However, it is not clear that including costs such as rising CEO salaries in the measure of variable costs would improve the interpretability of markup estimates.

Figure 1: *Data series of income inequality, firm entry and firm markups for eight OECD countries*

Another criticism of De Loecker et al. (2020) markup estimates relates to the heterogeneity of input elasticity across firms. When such heterogeneity is not taken into account, the production approach to markup estimation yields different estimates depending on the variable input used for estimation. Raval (2020) demonstrates this point by testing the similarity of markup estimates obtained using material and labour variable inputs. He finds that estimates generated with material and labour input are negatively correlated. He develops an estimator to model input elasticity heterogeneity across firms and find that markups thus generated are positively correlated in the cross-section and time dimensions.

Finally, Traina (2018) notes that estimated markups are positively correlated with firm size. Since public firms tend to be larger than the average firm (Davis et al., 2007), a measure of markups limited to public firms would tend to be biased upward.

The literature on markup estimation is evolving rapidly and improved estimates are expected to be available in the years to come. That exercise is beyond the scope of this paper; hence, I use the markup estimates that are currently available for the selected countries.

2.2 Methodology

Consider the model

$$y_{it} = \eta_i + \delta_i t + \beta_1 y_{i,t-1} + ... + \beta_p y_{i,t-p} + \epsilon_{it}, \quad t \in (t_{0i}, ..., T_i) \tag{1}$$

where y_{it} is a 4×1 vector for country $i \in (1, ..., N = 8)$ at time $t \in (t_{0i}, ..., T_i)$, p is the number of lags included in the model, T_i is the series length for country i and $\epsilon_{it} \sim iid(0, \Sigma_\epsilon)$. In a vectorial notation, the model can be expressed as

$$Y_i = \boldsymbol{\eta}_i \iota + \boldsymbol{\delta}_i \tau + B(L) Y_i + \mathcal{E}_i, \tag{2}$$

where Y_i is a $4 \times T_i$ vector of dependent variables, ι is a $4 \times T_i$ matrix of ones, τ is a matrix with a trend on each row and $B(L) = \beta_1 L + ... + \beta_p L^p$.

Bias Correction

Estimation presents two challenges: the bias induced by the shortness of the panel and the inclusion of country-specific trends that rules out the use of some existing bias alleviation strategies. It is well known that ordinary least squares (OLS) and maximum likelihood (ML) estimators of autoregressive models are biased towards stationarity in small sample (Hurwicz, 1950). With a small cross-section of $(N = 8)$ countries, I treat the cross-section dimension as fixed and consider the asymptotic properties of the estimator in the time dimension. I use a bootstrap-based bias correction approach proposed by MacKinnon and Smith (1998) and apply it in the context of dynamic panel estimation.

A number of common panel estimators are not suitable for this problem. The presence of a trend, however, makes the commonly used general method of moments (GMM) estimators unsuitable as they rely on differencing fixed-effects away which would not deal with the trend. The averaging estimator would provide an improvement compared to the OLS estimator, but the sample size is too short to estimate β separately for each country within a reasonable confidence interval. Sims (2000) suggests an interesting solution to reduce the bias, which is translated in practice by Alvarez and Arellano (2021) with a random effect unconditional maximum likelihood estimator (RML). As argued by Bai (2013), the RML estimator has a fixed effect interpretation, however it is not clear how this generalizes to the inclusion of a country-specific trend, and I am not willing to assume that the trend terms for each countries are drawn from the same distribution.

Some authors present analytical bias-correction methods (Phillips and Sul, 2007; Kiviet, 1995). However, these are developed for much simpler models and calculating the OLS bias analytically for (1), although theoretically possible, is beyond the scope of this paper.

I estimate (1) by maximum likelihood, conditioning on initial values of the dependent variables and followed the iterative bias correction method suggested by MacKinnon and Smith (1998). To my knowledge, this is the first attempt at applying a bootstrap bias-correction method to the estimation of a dynamic panel model. I provide simulation results below quantifying the remaining bias and I compare it to that of the uncorrected ML estimator. The bias correction algorithm and simulation results are provided in Appendix A.1.

Model Specification

Lag selection is determined by minimization of the Bayesian-Schwarz information criterion (BIC) conditional on the absence of significant serial correlation among residuals at a conservative level of 0.15. A portmanteau test for residual autocorrelation was performed separately for each of the eight countries in the unbalanced panel and a bootstrap-based procedure was applied to adjust for multiple testing. The procedure is described in detail in Appendix A.2. The set of selected lags is $\mathcal{L} = \{1, 5\}$.

Stationarity is established by testing for the cointegration rank of the system in its error correction representation. From the results of the Johansen cointegration tests detailed in Appendix A.3, I conclude that the system has full rank and is thus stationary, which is in line with economic theory in the long run since all variables are bounded rates.

2.3 Response Functions

In the theoretical model, presented below (Section 3), an increase in wealth inequality caused an increase in market power, via a decline in entrepreneurial creation of new firms. Such responses are observed empirically. I present hereinafter how firm markup responds to an inequality shock and how much of this response is propagated through the firm entry channel. I also discuss some evidence in support of the converse relation whereby inequality would also respond positively to a shock on firm markup.

I take an agnostic view as to the driver of inequality and define an inequality shock as any inequality variation that is not otherwise explained by the autoregressive model (1). Whether such variations are due to financial market shocks, skill biased technological changes or other economic driver may be important, but the drivers of inequality are left outside of the regression to make best use of the few degrees of freedom allowed by the short dataset.

The only control included is a series of interest rates. It is not unlikely that the results be driven by factors outside of the econometric model and results below suggesting causality should thus be interpreted cautiously. Nonetheless, the highlighted responses demonstrate the existence of some interesting regularities in the comovements of household inequalities and firm markups.

Identifying Assumptions

Additional structure is added to attribute an economic interpretation to the residuals of (1). Short-run restriction are provided by right multiplying the model by a matrix of contemporaneous relations A_0. Let

$$A_0^{-1} y_{it} = \gamma d_t + F(L) y_{it} + \nu_{it}, \quad t \in (t_{0i}, ..., T_i), \tag{3}$$

where $\gamma d_t = A_0^{-1}(\eta_i + \delta_i t)$, $F(L) = A_0^{-1} B(L)$ and $\nu_{it} = A_0^{-1} \epsilon_{it}$ is a $m \times 1$ vector of contemporaneously uncorrelated shocks. After normalizing with $I = E(\nu\nu')$ such that structural shocks are expressed relative to the standard deviations of the corresponding variables, A_0 is recovered using

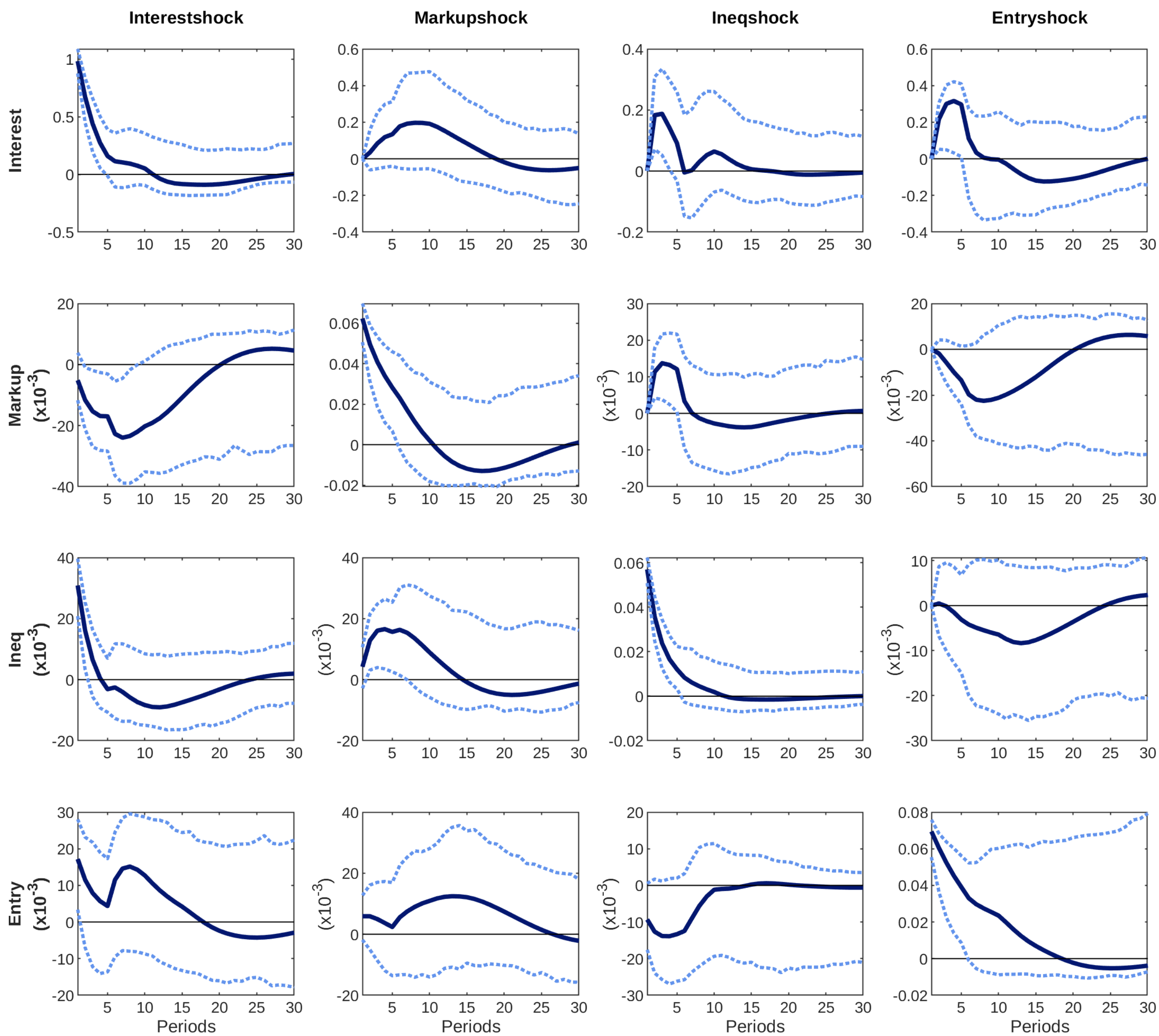
$$\Omega \equiv E(\epsilon\epsilon') = E(A_0 \nu\nu' A_0')$$

and

$$\hat{\Omega} = \frac{1}{T_i - 1} \sum_{t=t_{0i}}^{T_i} \epsilon_{it}\epsilon_{it}' = A_0 A_0'. \tag{4}$$

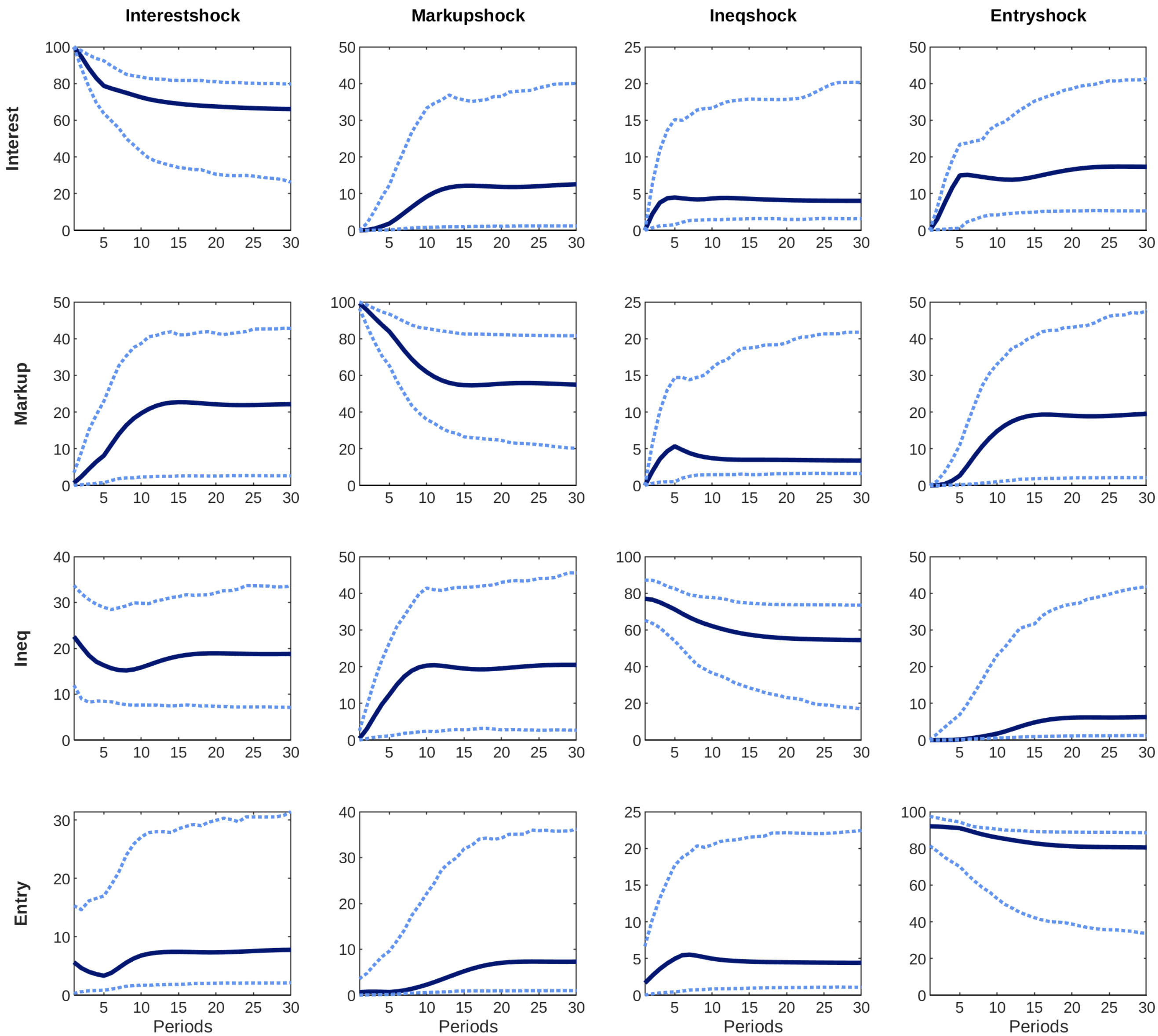
Three additional identifying restrictions are necessary to solve for a unique matrix of contemporaneous relations. They rely on the following three economic assumptions. First, I assume that the interest rate policy is contemporaneously unaffected by the other variables. It is not uncommon to model the interest rate as an exogenous policy variable. Here, I allow the interest rate to be endogenous and only restrict other variables contemporaneous effect on the interest rate. Second, I assume that neither the income share of the 99th percentile of households nor firm entry affect firms aggregate price-cost markup contemporaneously. To justify this assumption, I rely on the idea that price adjustments take time. Eichenbaum and Fisher (2007) estimate that firms readjust prices every 2.3 to 3 quarters. The third identifying assumption exploits the low contemporaneous correlation of 0.0053 between residuals of inequality and firm entry. I assume that firm entry does not affect income inequality contemporaneously.

Figure 2: *Response of variables (rows) to a positive shock (column)*



Note: Units are percentages for all variables. 85% confidence intervals computed by residual bootstrap with 499 repetitions.

Figure 3: *Share of the variance of variables forecast error (rows) attributed to each shock (columns)*



Note: 85% confidence intervals are computed by residual bootstrap with 499 repetitions.

Response of Firms Markups to Wealth Concentration Shocks

Figure 2 presents impulse response functions (IRF) for the four variables (rows) responding to corresponding shocks (columns). Three observations are in order. First, the response of inequality to firms markup supports the common idea that firms markup contributes to household income inequality as suggested by Furman and Orszag (2015), Stiglitz (2017). A second observation worth noting is that firms markup also seems to respond positively to variations in inequality. Figure 3 quantifies the contribution of inequality shocks to the variance of firms markups forecast error. Those two observations together suggest that a positive loop between inequality and markup may increase the persistence of markup shocks. I explore this possibility below by decomposing markup's persistence into the contribution from different channels. The third observation I wish to point to is the combination of the negative response of firm entry to an inequality shock and the negative response of markup to firm entry. These two IRFs combined suggest the existence of a firm entry channel through which inequality may affect aggregate firm markup.

Response of Firm Markups to Income Concentration Shocks via Firm Entry

I am interested in quantifying the contribution of the firm entry channel to the relationship between inequality and firm markup. For that purpose, I compute the response of markup to an inequality shock that propagates via the firm entry channel. The method is equivalent to comparing the relevant IRF with a counterfactual one where the firm entry channel would be shutoff as in Sims and Zha (2006), Bernanke et al. (1997) and Kilian and Lewis (2011). The method I propose is equivalent to Sims and Zha (2006)'s shutoff method and has the advantage of being very simple to compute from quantities readily available when performing IRF analysis. I offer a proof in the appendix.

Proposition 1 (IRF Channel Contribution) *Consider a $n \times 1$ stationary variable $X_t = (X_{qt}, X_{jt}, X_{it}, ...)'$. Let $\Phi_{qi}(L)$ be the power series describing the response of variable X_{qt} to a shock on variable X_{it}. Let $\bar{\Phi}_{qij}(L)$ be the power series describing the response of variable X_{qt} to a shock X_{it} transiting through channel $X_{jt}j$. Then,*

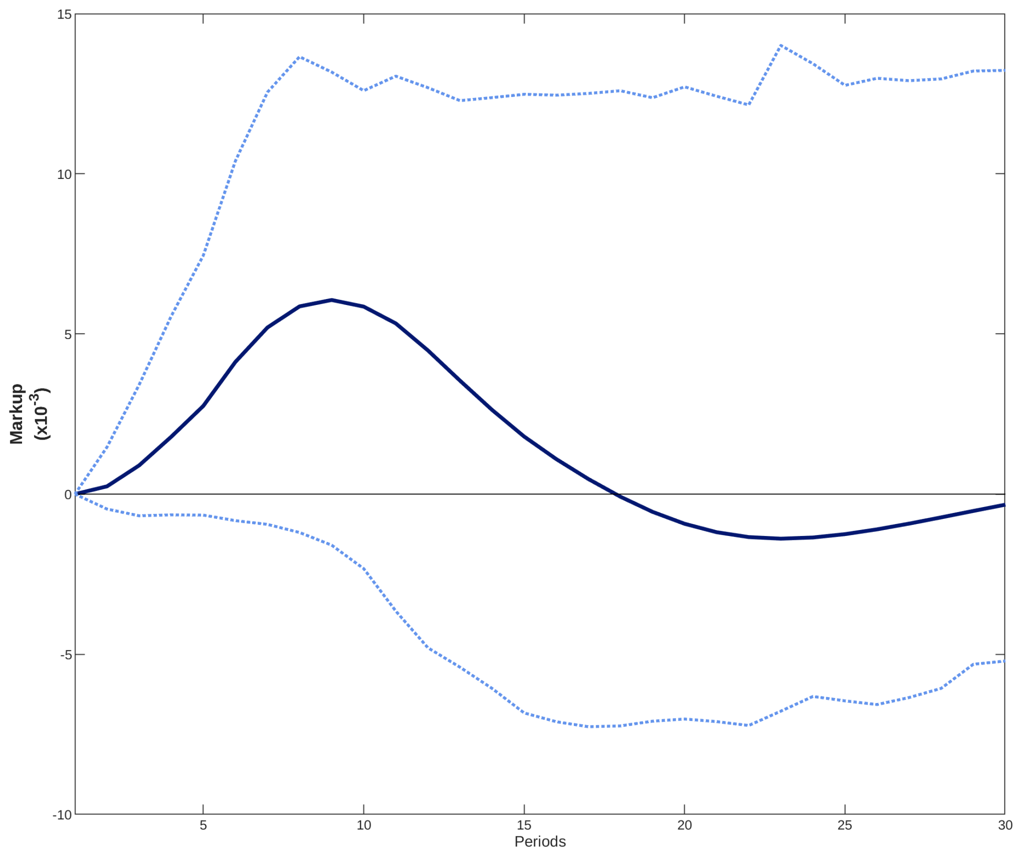
1. *$\bar{\Phi}_{qij}(L)$ is given by*

$$\Phi_{qj}(L) \times \Phi_{ji}(L) = \Phi_{jj}(L) \times \bar{\Phi}_{qij}(L); \tag{5}$$

2. *$\bar{\Phi}_{qij}(L)$ is equivalent to the channel-specific response obtained using Sims and Zha (2006)'s channel shutoff method.*

Using Proposition 1, the contribution of firm entry to the response of markup to an inequality shock is readily computed by convoluting the power series associated with the responses of markup to firm entry and of firm entry to inequality, truncating the obtained power series and dividing by the truncated power series describing firm entry's persistence to its own shocks. The method does not require additional computation when IRFs $\Phi_{qj}(L)$, $\Phi_{ji}(L)$, and $\Phi_{jj}(L)$ are readily available. The intuition behind it is that firm entry's persistence is taken into account in both the response of markup to firm entry and in the response of firm entry to inequality, thus the polynomial division by the truncated persistence of firm entry. If the truncation preserves a long enough horizon, this procedure is equivalent to the channel-shutoff method as demonstrated in Appendix C.

Figure 4: *Effect of inequality on markup through the firm entry channel*



Note: 85% confidence intervals are computed by residual bootstrap with 499 repetitions.

3 Theoretical Model

The following framework extends the step-by-step innovation model of Akcigit and Ates (2019) to include endogenous firm entry resulting from entrepreneurial effort conducted by households who differ only in their wealth. Akcigit and Ates (2019) formulation of the firm-side R&D problem is appealing in that it includes an endogenous markup measure and all the main variables can be expressed as a linear function of the total output Y_t. The latter allows for a simple normalization under which all main values of interest are either constant or stationary on a balanced growth path (BGP).

The model comprises a unit measure of households who consume, save (borrow), work for a fixed wage and have the possibility to engage in entrepreneurial activity at some cost. Households jointly solve their consumption and entrepreneurial investment decisions, reflecting the idea that entrepreneurial firms access to credit is often linked to their owner's wealth (Davis and Haltiwanger, 2019)[3] and that they are more risk averse than other firms because entrepreneurs cannot diversify away their idiosyncratic risk (Caggese, 2012; Quadrini, 2000).

A representative final goods producer takes a continuum of intermediate goods as input and aggregates them into a unique consumption good to be enjoyed by households. Intermediate goods from each industry denoted $j \in [0, 1]$ is produced according to a patent held by the leading firm in that industry. Each industry j is either dominated by a monopolist firm or two toe-to-toe firms competing for leadership. The share μ of monopolistic industries depends on the rate of firm entry $\mathcal{Z}$, which is driven by households entrepreneurial efforts.

Two key features are worth noting. First, both firm markup and the distribution of household wealth are endogenously determined. This allows to discuss the effect of firm markup on wealth inequality and vice-versa and compare such effects when different underlying drivers are at play.

A second interesting feature of the model is the joint optimization by households of their consumption together with a quantitative entrepreneurial investment decision. A body of literature on entrepreneurial idiosyncratic risks already incorporates such joint optimization problem, but it is seldom seen in the Schumpeterian endogenous growth literature. García-Peñalosa and Wen (2008) were perhaps precursors in this regard but only consider a partial equilibrium framework.

The model is described in detail below. I first specify households preferences, consumption goods production and the intermediate goods monopolists' profits. In the second subsection, I describe technology firms competition for leadership. Then, I present households

[3]Davis and Haltiwanger (2019) find that young firm activity shares move strongly with local economic condition and local house price growth. The authors find statistical evidence to the effect that local house prices disproportionately affect young firms, suggesting the existence of a linkage between firm credit and owners' wealth. This view is also shared by Quadrini (2000) and is coherent with (Fazzari et al., 1987; Gertler and Gilchrist, 1994; Gilchrist and Himmelberg, 1995).

problem and their entrepreneurial activity decision.

3.1 Consumer Preferences, Technology Protection and Firms Profit

Consumer preferences, final goods aggregator's problem and firms profit are as proposed by Akcigit and Ates (2019). I restate them here for convenience. The firms problem departs from previous literature with the introduction of the entrepreneurially driven rate of firm entry $\mathcal{Z}$.

Preferences

Consider a world where a mass M of households denoted i maximize their lifetime utility

$$U_{it} = \int_t^\infty e^{-\rho(\tau - t)} u(c_{i\tau}) d\tau, \tag{6}$$

where c_{it} is their consumption at time t, $\rho > 0$ is the discount rate and constant relative risk aversion (CRRA) preferences are expressed with the utility function

$$u(c) = \begin{cases} \log(c), & \gamma = 1 \\ \frac{c^{1-\gamma}-1}{1-\gamma}, & 0 < \gamma \neq 1. \end{cases}$$

Production

Households consume a unique final good produced according to the following Cobb-Douglas production function

$$\log(Y_t) = \int_0^1 \log(y_{jt}) dj, \tag{7}$$

where y_{jt} is the output of intermediate goods of variety $j \in (0, 1)$. In this setting, the role of the final goods producer is merely to externalize the constant elasticity of substitution (CES) aggregation from households utility function to simplify households problem.

Intermediate goods j are produced by industry j's leading firm(s) with an amount l_{jt}^P of labour working with productivity q_{jt} according to the linear production function

$$y_{jt} = q_{jt} l_{jt}^P.$$

Industry j's productivity is given by $q_{jt} = \lambda^{k_{jt}}$, with $k_{jt} \in \mathbb{N}_+$ being the number of R&D improvements until time t, and λ being the productivity step between subsequent improvements.

Profits

The demand for an intermediate good j priced p_{jt} is obtained from the first order condition of the representative final good producer's problem and is given by

$$y_{jt} = \frac{Y_t}{p_{jt}}. \tag{8}$$

The production cost of one unit of y_{jt} is

$$MC_{jt} = \frac{w_t}{q_{jt}}, \tag{9}$$

where w_t is the equilibrium wage. When, in an industry j, the firm enjoys a monopoly on the more advanced technology for the production of y_{jt}, they set their selling price p_{jt} at the marginal cost of potential competitors who would only have access to technology $\lambda^{k_{jt}-1}$. However, in industries where the leadership is shared by two toe-to-toe leaders, these leaders compete in price and the price is given by their own marginal costs. Thus,

$$p_{jt} = \begin{cases} \frac{w_t}{\lambda^{(k_{jt}-1)}} & , m_{jt} = 1 \\ \frac{w_t}{\lambda^{(k_{jt})}} & , m_{jt} = 0, \end{cases} \tag{10}$$

$$= \frac{w_t}{\lambda^{(k_{jt}-m_{jt})}}, \tag{11}$$

where $m_{jt} \in \{0, 1\}$ indicates the number of innovation steps separating the leading firm and it's closest competitor at time t. Substituting the price p_{jt} into the demand function for y_{jt} (8) yields the equilibrium intermediate goods quantities.

$$y_{jt} = \frac{Y_t}{w_t} \lambda^{(k_{jt}-m_{jt})}. \tag{12}$$

When $m_{jt} = 0$, toe-to-toe leaders compete à la Bertrand and no profit is earned. When $m_{jt} = 1$, the profit earned by the monopolist in industry j is

$$\begin{aligned} \pi_{jt} &= \left(p_{jt} - MC_{jt}\right) y_{jt} \\ &= \left(w_t \lambda^{-(kjt-1)} - w_t \lambda^{-k_{jt}}\right) \frac{Y_t}{w_t} \lambda^{(k_{jt}-1)} \\ &= \left(1 - \frac{1}{\lambda}\right) Y_t, \end{aligned} \tag{13}$$

and their markup is given by

$$\mathcal{M}_{jt} = \frac{p_{jt}}{MC_{jt}} - 1 = \lambda - 1.$$

Note that equilibrium profit π_{jt} and markup $\mathcal{M}_{jt}$ are the same for all product lines and will thus be denoted π_t and $\mathcal{M}_t$ henceforth.

Competing Firms: A Step-by-step Innovation Process

Every industry j can be in one of two states $m_j \in \{0, 1\}$. When $m_j = 1$, the unique leading firm is enjoying a monopoly on technology good j and a profit flow π_t. When an entrepreneur catches up with the leader, the leadership position is shared between the incumbent leader and a new firm created by the entrepreneur. The entering firm is sold by the entrepreneur to risk neutral investors who can perfectly diversify their exposure to idiosyncratic risk and who compete to obtain a share of the new firm. After entry, $m_j = 0$ as the two leading firms are toe-to-toe. Since firms compete à la Bertrand neither firm makes a positive profit until one of them innovates further.

Innovation success for toe-to-toe firms in industry j follows a Poisson distribution with arrival rate $\zeta(l^f_{jt})$ given by the concave innovation function

$$\zeta(l^f_{jt}) = \phi_f (l^f_{jt})^\eta, \quad 0 < \eta < 1, \tag{14}$$

where ϕ_f is an innovation scale parameter to be calibrated and l^f_{jt} is the R& D labour employed at time t by firms of line j. The decreasing marginal return on innovation investment results from simpler and cheaper ideas being explored first (Cavenaile et al., 2020).

Only the most advanced technology is protected and the greatest lead a firm can have on any competitor is one step. For simplicity, I assume that when an entrepreneur successfully innovates in an industry where leadership is already shared, the newly created firm replaces one of the two leaders with equal probability. This assumptions allows for a uniform new firm value irrespective of the industry they enter in. A toe-to-toe leading firm that becomes a laggard stops its activities.

Firms Value Functions

Let $\mathcal{V}_{m_j,t}$ be the value of an existing firm in any industry j. On a balanced growth path (BGP), the interest rate r and the aggregate rate of entrepreneurial innovation $\mathcal{Z}$ are constant. In the unleveled industry, dominated by a monopolist, we have

$$r\mathcal{V}_{1,t} - \dot{\mathcal{V}}_{1,t} = \left(1 - \frac{1}{\lambda}\right) Y_t + \mathcal{Z}\left(\mathcal{V}_{0,t} - \mathcal{V}_{1,t}\right).$$

In the leveled industry, the two toe-to-toe leaders have a value given by

$$r\mathcal{V}_{0,t} - \dot{\mathcal{V}}_{0,t} = \max_{l^f_t} \left\{ -w\, l^f_t + \zeta(l^f_t)\left(\mathcal{V}_{1,t} - \mathcal{V}_{0,t}\right) - \left(\zeta({l^f_t}^*) + \tfrac{1}{2}\mathcal{Z}\right)\mathcal{V}_{0,t} \right\},$$

where $l_t^{f^*} = l_t^f$ is the competitor's innovation effort decision. Expressing values relative to Y_t and defining $\nu_{m_j} \equiv \mathcal{V}_{m_j,t}/Y_t$, normalized firms BGP value functions become

$$\tilde{r}\nu_1 = \left(1 - \tfrac{1}{\lambda}\right) + \mathcal{Z}\left(\nu_0 - \nu_1\right), \tag{15}$$

$$\tilde{r}\nu_0 = \max_{l^f}\left\{-\tilde{w}\, l^f + \zeta(l^f)\left(\nu_1 - \nu_0\right) - \left(\zeta(l^{f^*}) + \tfrac{1}{2}\mathcal{Z}\right)\nu_0\right\}. \tag{16}$$

The interior solution to the toe-to-toe leaders' problem solves the first order condition

$$\begin{aligned}
\frac{\partial \tilde{r}\nu_0}{\partial l^f} &= -\,\tilde{w} + \eta\phi_f (l^f)^{\eta-1}(\nu_1 - \nu_0) = 0 \\
(l^f)^{\eta-1} &= \frac{\tilde{w}}{\eta\phi_f(\nu_1 - \nu_0)} \\
l^{f^*} &= \left(\frac{\eta\phi_f(\nu_1 - \nu_0)}{\tilde{w}}\right)^{\frac{1}{1-\eta}}.
\end{aligned} \tag{17}$$

The firm's problem always has an interior solution because the marginal value of R&D is infinite at the origin.

Market Power

Following Akcigit and Ates (2019), market power is proxied by the share $\mu_t \in [0, 1]$ of industries led by a monopolist firm. The share of unleveled industries is determined by the innovation rate of toe-to-toe leaders in leveled industries (ζ) and the innovation rate of entrepreneurs ($\mathcal{Z}$) with

$$\dot{\mu}_t = 2(1 - \mu_t)\zeta(l_t^f) - \mu_t \mathcal{Z}_t. \tag{18}$$

On a BGP, $\dot{\mu} = 0$ and the stationary share of unleveled industries is

$$\mu = \frac{\zeta(l^f)}{\zeta(l^f) + \frac{1}{2}\mathcal{Z}}. \tag{19}$$

The average firm markup $\mathcal{M}$ is given by

$$\mathcal{M} = \mu(\lambda - 1) = \frac{\zeta(l^f)}{\zeta(l^f) + \frac{1}{2}\mathcal{Z}}(\lambda - 1). \tag{20}$$

Equation (19) makes explicit the first order effect of entrepreneurial activity on market power in this model. The share of unleveled industries is decreasing in $\mathcal{Z}$. As entrepreneurial activity increases, firms competitiveness is stimulated. The next subsection discusses households choice of entrepreneurial effort.

3.2 Household Decisions and Entrepreneurial Activity

For entrepreneurs, the reward from a successful entry is given by ν_{0t}, the value of a toe-to-toe firm. Upon success, entrepreneurs sell their new project at the fair price $\Pi_t = \nu_{0t}$. Their Poisson rate of success $z(l^e_{it})$ follows the same functional form as that of the firm with

$$z(l^e_{it}) = \phi_e (l^e_{it})^\eta, \quad 0 < \eta < 1, \tag{21}$$

where ϕ_e is an entrepreneurial scale parameter and concavity parameter η is the same for firms innovation function and households entrepreneurial activity function $(z(l^e_{it}))$.

On a balanced growth path (BGP), households' problem is given by this properly normalized Hamilton-Jacobi-Bellman (HJB) equation where all pecuniary variables are expressed relative to output Y with $\tilde{c}_{it} \equiv \frac{c_{it}}{Y_t}$, $\tilde{a}_{it} \equiv \frac{a_{it}}{Y_t}$, $\tilde{w}_t = \frac{w_t}{Y_t}$ and $\tilde{r}_t \equiv r_t - g_t$. On a BGP, normalized variables are constant and the time subscript is omitted to ease notation. The HJB equation is derived in Appendix B. Households solve

$$\psi v(\tilde{a}) = \max_{\tilde{c}, l^e} u(\tilde{c}) + v'(\tilde{a})\dot{\tilde{a}} + z(l^e)\left(v(\tilde{a} + \Pi) - v(\tilde{a})\right) \tag{22}$$

subject to

$$\dot{\tilde{a}} = \tilde{r}\tilde{a} + \tilde{w}(1 - l^e) - \tilde{c}, \tag{23}$$

$$\tilde{a} \geq \underline{\text{a}}, \tag{24}$$

where $\psi = \rho - (1 - \gamma)g$ is the discount rate for the normalized value function and $\underline{\text{a}}$ is the normalized borrowing constraint. Households normalized flow of value, in terms of utility, is comprised of their contemporaneous utility, the value of the change in savings and the expected value of their entrepreneurial endeavor, driven by their Poisson rate of entrepreneurial success $z(l^e)$ and the startup value $\Pi = \nu_0$ obtained upon success.

Households first order conditions are given by

$$v'(\tilde{a}) = \tilde{c}^{-\gamma} \tag{25}$$

$$l^e = \left(\eta\phi_e \frac{[v(\tilde{a} + \Pi) - v(\tilde{a})]}{\tilde{w}v'(\tilde{a})}\right)^{\frac{1}{1-\eta}}. \tag{26}$$

Aggregate Savings and Wealth Distribution

Relative household wealth $\tilde{a}_t$ evolves according to the savings function $s_t(\tilde{a}_t)$ given by (23) as

$$d\tilde{a}_t = s_t(\tilde{a}_t)dt, \tag{27}$$

where $s_t(\tilde{a}_t) = \tilde{r}\tilde{a} + \tilde{w}(1 - l^e) - \tilde{c}$. The state of the economy is determined by quality levels $\{k_{jt}\}$ and the relative wealth distribution $f_t(\tilde{a})$. We assume that an entrepreneur's project is randomly assigned to a line j upon success. Consider the cumulative distribution function

$$F_t(\tilde{a}) = \Pr(\tilde{a}_t \leq \tilde{a}|t)$$

with

$$F_t(\underline{a}) = 0; \; \lim_{\tilde{a}\to\infty} F_t(\tilde{a}) = 1; \text{ and } f_t(\tilde{a}) \equiv F_t'(\tilde{a}) \; \forall t.$$

The time-independent Kolmogorov forward equation for a BGP stationary wealth distribution is

$$0 = -\frac{\partial}{\partial \tilde{a}}\left[(\tilde{r}\tilde{a} + \tilde{w}(1 - l^e(\tilde{a})) - \tilde{c}(\tilde{a}))f(\tilde{a})\right] - z(l^e(\tilde{a}))f(\tilde{a}) + z(l^e(\tilde{a} - \tilde{\Pi}))f(\tilde{a} - \tilde{\Pi}). \quad (28)$$

This result is derived in Appendix B.

3.3 Aggregate Wage and Output

Akcigit and Ates (2019) define the aggregate productivity index Q_t, which is useful in deriving the aggregate wage and output.

$$Q_t = e^{\int_0^1 \log q_{jt} dj} \quad (29)$$

$$= e^{\int_0^1 k_{jt} \log(\lambda) dj} \quad (30)$$

Aggregate Wage

Combining the final good production function (7) and the equilibrium intermediate goods quantity (12), the equilibrium wage is derived as follows

$$Y_t = e^{\int_0^1 \log y_{jt} dj}$$

$$Y_t = e^{\int_0^1 (k_{jt} - m_{jt}) \log(\lambda) dj} \cdot \frac{Y_t}{w_t}$$

$$w_t = e^{\int_0^1 k_{jt} \log(\lambda) dj - \mu_t \int_0^1 \log(\lambda) dj}$$

$$w_t = \frac{Q_t}{\lambda^{\mu_t}}. \quad (31)$$

The real labour wage $\tilde{w} \equiv \frac{w_t}{Y_t}$ can be obtained from the intermediate goods labour $\int_0^1 l_{jt}^P dj = L_t^P$ and the share of profits. Total monopolist profits are given by

$$\int_0^1 \pi_t dj = \mu_t \pi_t = \mu_t \left(1 - \tfrac{1}{\lambda}\right) Y_t$$

and total wages for intermediate goods production are

$$\int_0^1 l_{jt}^P w_t dj = L_t^P w_t.$$

Attributing income from intermediate goods sales to either wages or profits solves for the wage

$$\begin{aligned}
Y_t =& L_t^P w_t + \mu_t \left(1 - \tfrac{1}{\lambda}\right) Y_t \\
\tilde{w} \equiv \frac{w_t}{Y_t} =& \frac{1}{L_t^P} \left(1 - \mu_t (1 - \tfrac{1}{\lambda})\right). \qquad (32)
\end{aligned}$$

In equilibrium, the labour market clears with

$$M = L_t^P + \int_0^M l_{it}^e di + 2 \int_0^1 m_{jt} l_{jt}^f dj. \qquad (33)$$

Aggregate Output

Combining expressions for the equilibrium wage (31) and the wage share of output (32) yields an expression for the aggregate output

$$\begin{aligned}
Y_t =& \frac{L_t^P w_t}{1 - \mu_t \left(1 - \frac{1}{\lambda}\right)} \\
Y_t =& \frac{L_t^P Q_t}{\lambda^{\mu_t} \left[1 - \mu_t \left(1 - \frac{1}{\lambda}\right)\right]}. \qquad (34)
\end{aligned}$$

3.4 Output Growth

On a BGP, aggregate firm entry rate $\mathcal{Z}$ is constant and the aggregate output growth rate g is derived as follows

$$\begin{aligned}
\log(Q_t) =& \int_0^1 \log(q_{jt}) dj \\
\log(Q_{t+\Delta t}) =& \int_0^1 (1-\mu) 2\zeta(l^f) \Delta t \log(\lambda q_{jt}) dj + \int_0^1 \Big(1 - (1-\mu) 2\zeta(l^f)\Delta t\Big) \log(q_{jt}) dj \\
\log(Q_{t+\Delta t}) - \log(Q_t) =& (1-\mu) 2\zeta(l^f) \Delta t \int_0^1 \log(\lambda) dj \\
\frac{\log(Q_{t+\Delta t}) - \log(Q_t)}{\Delta t} =& (1-\mu) 2\zeta(l^f) \log(\lambda) \\
g = \frac{\dot{Q}_t}{Q_t} =& (1-\mu) 2\zeta(l^f) \log(\lambda). \qquad (35)
\end{aligned}$$

It is clear from (35) that BGP growth is increasing with the share of leveled firms $(1-\mu)$ and firms R&D effort (l^f).

3.5 Balanced Growth Path Equilibrium

On a BGP, successful entrepreneurs sell their project at a fair price and receive the value of a toe-to-toe firm ν_0.

At all times the financial assets market clears such that aggregate wealth is equal to existing public firms equity. Total equity $\mathcal{E}$ is given by the value of existing firms.

$$\mathcal{E} = \mu\nu_1 + 2(1-\mu)\nu_0. \tag{36}$$

In equilibrium, the interest rate is such that

$$\tilde{A} = M\int_{\underline{a}}^{\bar{a}} \alpha f(\alpha) d\alpha = \mathcal{E}. \tag{37}$$

In this Schumpeterian endogenous growth model with heterogeneous entrepreneurs differing in their wealth, a BGP Markov perfect equilibrium is an allocation $y_{jt}, p_{jt}, C_t, Y_t, L$, a value function $v(\tilde{a})$ and associated policy functions $\tilde{c}(\tilde{a})$ and $l^e(\tilde{a})$ for entrepreneur households, value functions $\{\nu_0, \nu_1\}$ and associated policy function l^f for firms, relative prices $\tilde{w}$, $\tilde{r}$, $\tilde{\Pi}$, and stationary distributions $f(\tilde{a})$ and μ such that

- Given relative prices $\tilde{w}$, $\tilde{r}$ and $\tilde{\Pi}$, policy functions $\tilde{c}(\tilde{a})$ and $l^e(\tilde{a})$ solve the household's problem and $v(\tilde{a})$ from equation (56) is the associated value function;
- Given the relative interest rate $\tilde{r}$, the wage $\tilde{w}$ and the firm entry rate $\mathcal{Z}$, the toe-to-toe firm's demand for R&D labour (l^f) solves the firm's R&D problem, and $\{\nu_0, \nu_1\}$ are the associated firms value functions;
- Relative wage $\tilde{w} \equiv \frac{w_t}{Y_t}$ clears the labour market at all times such that (33) is satisfied;
- Relative interest rate $\tilde{r} \equiv r - g$ clears the financial market at every instant, satisfying equation (37);
- The aggregate entry rate satisfies $\mathcal{Z} = M\int_{\underline{a}}^{\bar{a}} z(l^e(\alpha)) f(\alpha) d\alpha$;
- The stationary distribution $f(\tilde{a})$ of household wealth satisfies (63);
- The stationary measure of leveled firms (μ) satisfies (19);
- Aggregate variables (C_t, Y_t, w_t) grow at a constant rate g given by (35);
- The resource constraint is satisfied at all times with

$$1 = \int_{\underline{a}}^{\infty} \tilde{c}(\alpha) f(\alpha) d\alpha.$$

4 Calibration

To ease interpretation of the results, the set of parameters $\{M, \eta, \gamma, \rho, \lambda, \phi_e, \phi_f, \underline{a}\}$ is calibrated with values of $\{M, \eta, \gamma\}$ determined from the relevant literature and values of $\{\rho, \lambda, \phi_e, \phi_f\}$ chosen to fit four moments from U.S. data: The real GDP per capita average growth rate between 1980 and 2008, the average equity return rate over the same period, the aggregate firms markup in 1980 estimated by De Loecker et al. (2020), and the 1980 firm entry rate compiled from local business registries and made available by the OECD. The borrowing constraint ($\underline{a}$) is chosen to be finite, but unbinding, so that households own budget constraint and risk aversion are their only effective constraints on borrowing.

The list of all calibrated parameters with their values is displayed in Table 1. The average number of households per firm (M) determines the value of a new firm relative to workers wage. Higher M yields a longer right tail of the wealth distribution as households earn a higher innovation reward, but they succeed less frequently. M is set such that the labor force employed in production (L^P) is equal to the average number of workers per firm obtained from the 2007 U.S. Census Bureau's Survey of Business Owners.

The innovation function curvature η is the same for both firms and entrepreneurs and it determines the return to scale of the innovation process. Kortum (1993) report estimates of η between 0.1 and 0.6. This parameter is central to the mechanics between the wealth distribution and firm entry. A low value of η implies that, from a planner's perspective, it is more efficient to have the innovation effort distributed equally among entrepreneurs. I use $\eta = 0.6$ to be conservative. Figure 5 compares the curvature of the share of labour hours invested in entrepreneurial R&D implied by η with the share of self-employed among employed households according to 1983's Survey of Consumer Finances (SCF). Points on the scatter plot represent weighted shares of self-employed within subsequent bins of 20 respondents sorted by their wealth. The pattern suggests a concave relationship between self-employment and wealth.[4] This provides an additional argument for a choice of η below unity.

Households risk aversion γ is a key component of the dynamics presented in this model. Since households are not allowed to default and no external borrowing constraint is imposed on them, households borrowing limitations are determined solely by their idiosyncratic wealth level and their risk aversion.[5] The higher the risk aversion, the more the wealth distribution affects aggregate entrepreneurial activity. Most widely accepted values for γ lie between 1 and 3 (Gandelman and Hernandez-Murillo, 2015). I use $\gamma = 2.5$.

[4]Repeating the exercice with the 2019 SCF survey yields the same concave relationship between self-employment and wealth for most of the wealth distribution except at the very top of the distribution where the pattern is more consistent with a convex relationship.

[5]As an extension, it could be interesting to incorporate default to the model and make the price of

Figure 5: *Entrepreneurial activity as a function wealth*

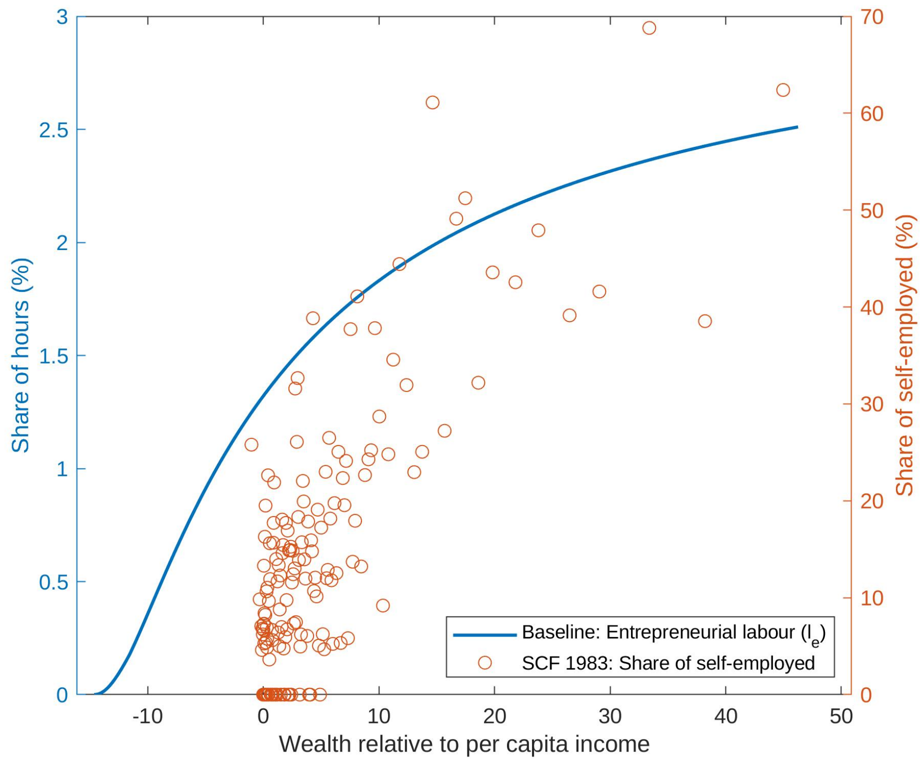


Table 1: List of parameter values

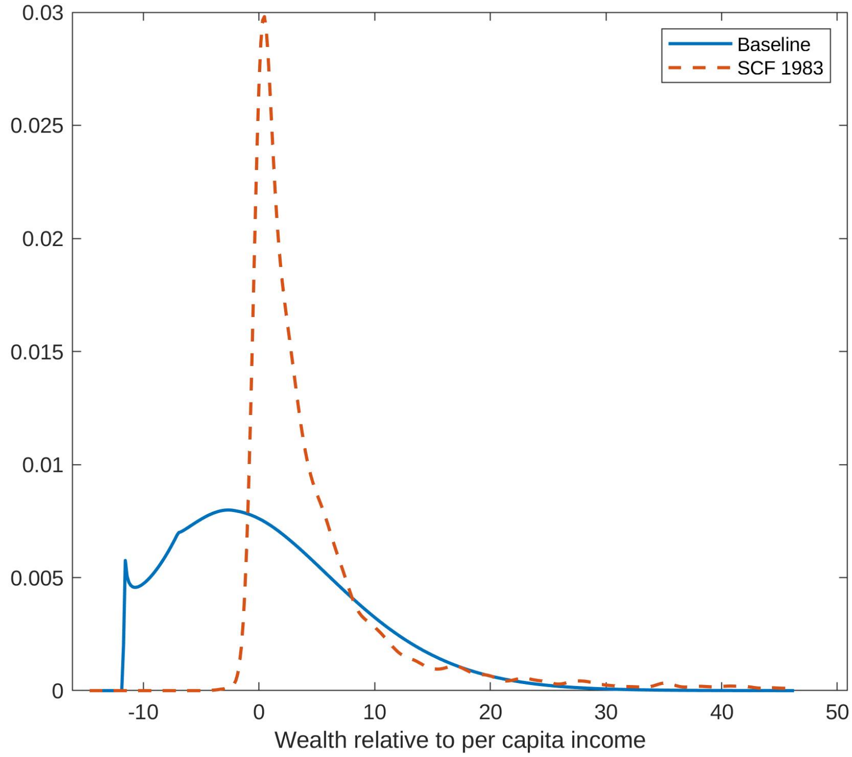
Externally Calibrated Parameters			**Internally Calibrated Parameters**		
	Description	Value		Description	Value
M	Households per firm	22	ρ	Time discount rate	0.04
η	R&D function concavity	0.6	λ	Innovation step	1.162
γ	Household risk aversion	2.5	ϕ_e	Entrepreneurs R&D productivity	0.0833
			ϕ_f	Firms R&D productivity	0.224

The remaining four parameters are selected to fit the model to four targeted moments. The time discount ρ is chosen to fit the BGP interest rate r to the average Standard & Poor 500 annual returns between 1980 and 2008, taking dividends and inflation into account. λ is selected to fit the BGP growth g to the real GDP per capita average annual growth rate in the U.S. between 1980 and 2008. Aggregate firm entry rate $\mathcal{Z}$ is driven by entrepreneurs R&D productivity ϕ_e. U.S. firm entry rate of 1980 is considered. It is computed from relevant business registries and made available by the OECD. Finally, firms R&D productivity ϕ_f is determined to fit the model's markup $\mathcal{M}$ to the aggregate firm price markup for the U.S. in 1980 estimated by De Loecker et al. (2020). Other estimates of markup exist. For instance, (Akcigit and Ates, 2019) use the average of the estimates given by De Loecker et al. (2020)

borrowing contingent on households idiosyncratic risk associated with their wealth level.

and Eggertsson et al. (2018). For the sake of the calibration exercise, I use De Loecker et al. (2020)'s estimate to be consistent with the data choice discussed above.

Figure 6: *Wealth distribution*



The resulting wealth distribution for the baseline model is displayed in Figure 6 and compared with the wealth distribution given by the SCF of 1983. The baseline model's wealth distribution features a long right tail. However, households accumulate more debt than what is empirically observed. This is due to the absence of an exterior borrowing constraint combined with the appeal to invest in entrepreneurial activity for a great potential reward.[6]

The theoretical model is offered as a proof of concept and has been kept as lean as possible to highlight the effect of the main dynamics between inequality, firm entry and market power. Of course, many components such as heterogeneity in ability, borrowing constraints and heterogeneity in returns could be added to better fit the data. The main purpose of the calibration exercise is to ease interpretation of the simulation results presented in the next section.

[6] An interesting extension would be to allow default and have households price of credit be modulated by their defaulting risk. This would shorten the left tail of the distribution without households massing around a natural borrowing limit.

Table 2: Targeted Moments

Moment	Model	Data	Source
Real GDP per capita growth rate	1.78%	1.676%	Federal Reserve Bank of Saint-Louis
Equity return rate	8.15%	8.05%	Average S&P 500 returns
Aggregate firm price markup	1.1474	1.1470	De Loecker et al. (2020)
Firm entry rate	13.047%	13.068%	OECD

5 Simulations

In the proposed framework, both the wealth distribution and the measure of firm markup are endogenous. This allows for a richer general equilibrium analysis. However the endogeneity of the wealth distribution makes the study of its effects more challenging. I detail, in this section, how I isolate effects transiting through the variation in the shape of the wealth distribution.

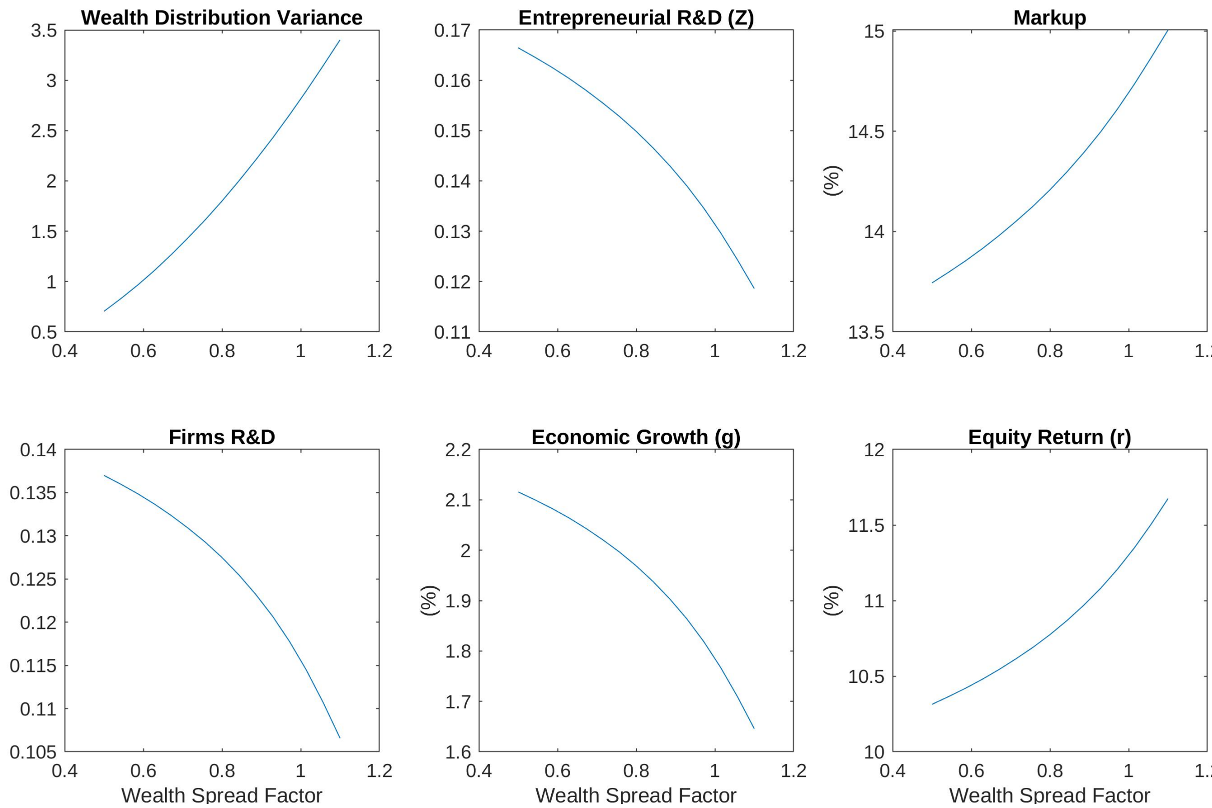
I offer two simulation exercises to gain insight into the effect of wealth concentration on macroeconomic variables of interest. First, a static comparative exercise displays the partial equilibrium consequence of a mean preserving spread in the wealth distribution on firm entry, markup, incumbent firms R&D and growth. I find that, with prices held fixed, a mean preserving spread in the wealth distribution reduces firm entry and increases market power. As a second exercise, I simulate the full equilibrium effect of variations of a selection of drivers of wealth inequality. I decompose the effect into an incentive (distortionary) effect and a redistributive effect. Different drivers of inequality change the shape of the wealth distribution differently and, as a consequence, the effects on macroeconomic variables of interest associated with a spread in the wealth distribution varies slightly depending on the underlying driver of inequality. In general, a concentration of wealth is associated with an increase in market power and a decrease in firm entry and growth around the calibrated benchmark.

Finally, I compare and discuss the effects of five fiscal policies (three taxes and two subsidies). I find, among other results, that a slight redistributive wealth tax reduces markup and enhances firm entry and growth.

5.1 Static Comparative

As a first simulation exercise, I investigate the partial equilibrium effects of a mean preserving spread of the wealth distribution on firm entry ($\mathcal{Z}$), firm markup ($\mathcal{M}$), aggregate firms R&D, output growth (g) and equity returns (r). I apply a uniform wealth spread factor ξ that multiplies each household's wealth deviation from the mean. The spread factor ranges from 0.5 to 1.1, where 1 corresponds to the calibrated BGP equilibrium wealth distribution. I limit the spread to 1.1 to avoid indebted households wealth to be pushed below the natural borrowing limit.

Figure 7: *Partial equilibrium effect of a mean preserving spread of the wealth distribution*



Note: The calibrated benchmark corresponds to the case where the wealth spread factor equals one.

Figure 7 presents the partial equilibrium effects as a function of the spread factor ξ. An increase in the variance of the wealth distribution causes a decrease in the firm entry rate ($\mathcal{Z}$). and an increase in the firm price markup. These findings are a result of the assumptions of a decreasing marginal return to entrepreneurial effort ($\eta < 1$) and the uninsurable nature of entrepreneurial risk. The empirical exercise discussed in Section 2 exhibits markup responses to inequality shocks that are consistent with this result.

The decrease in entrepreneurial expenditure and firm entry as a consequence of the increasing inequality results in a reduction in firm competitiveness and a higher share of monopolist industries. The higher wealth concentration thus yields reduced firm R&D and growth.

The static comparative exercise does not take price adjustments into account. Equilibrium effects of changes in the wealth distribution are discussed next.

5.2 Drivers of Inequality and Equilibrium Effects

In this section, I compute full equilibrium effects of changes in wealth inequality and I investigate how the elasticity of macroeconomic variables to inequality variations differ according to the drivers of inequality. To compute that elasticity, I isolate effects transiting via the shape of the wealth distribution by comparing the total effect of a permanent change of an exogenous driver with the effect obtained when the shape of the wealth distribution is unchanged. To calculate the latter, I let the mean of the distribution adjust to the point of zero aggregate savings[7] while imposing a series of continuous "surprise" lump sum transfers that agents do not incorporate in their policy functions so that the shape of the distribution remains fixed.

Four exogenous drivers for the wealth distribution are considered. I investigate the effect of changes in the innovation step size (λ), the R&D function curvature (η), entrepreneurs R&D productivity (ϕ_e) and a wealth tax (τ_W) with transfers. Households being skill-wise homogeneous in the current framework, it is not possible to include skill-biased technical change among the drivers of inequality.

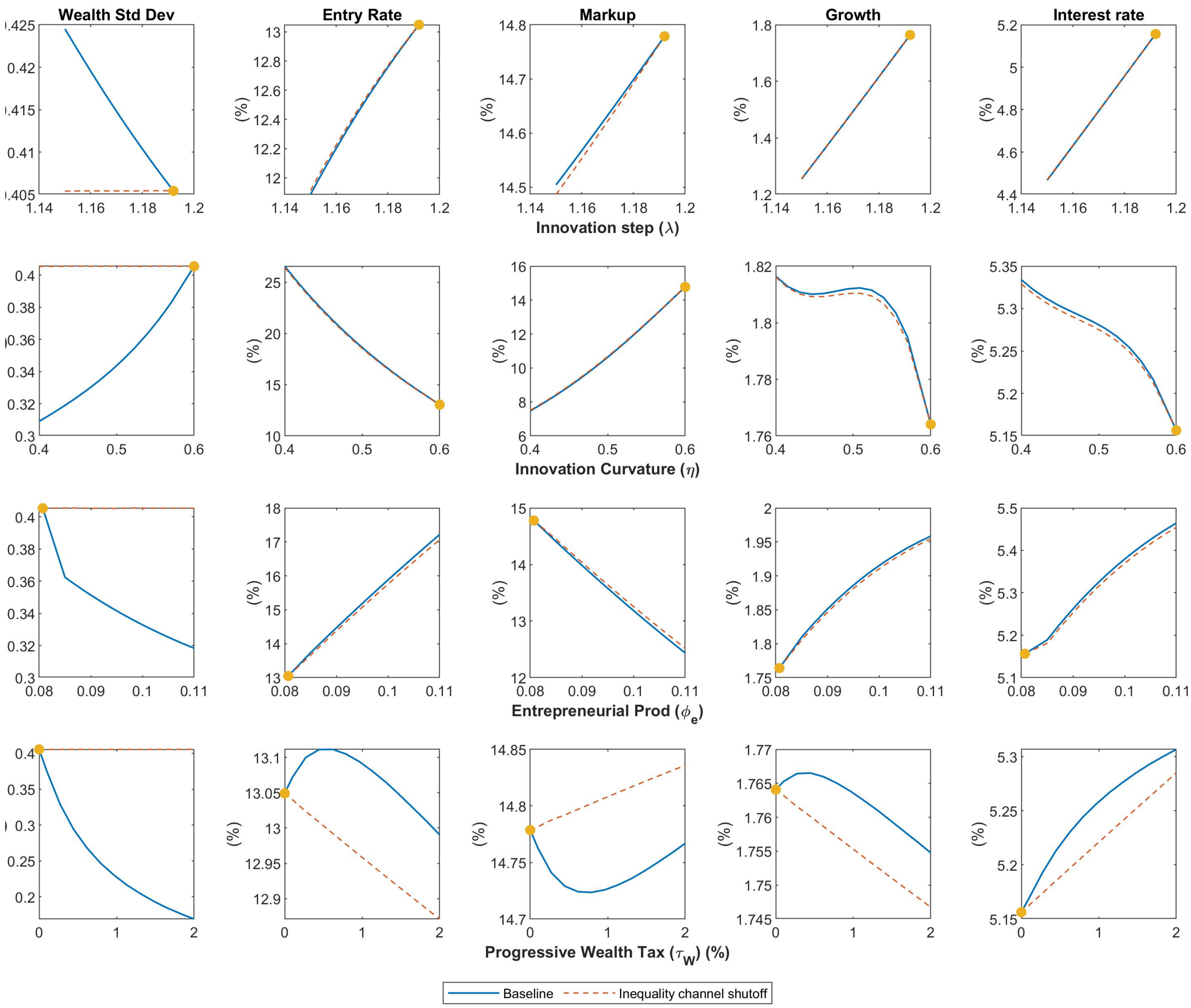
Bloom et al. (2020) suggest that "ideas are getting harder to find" and provide the example of computer chips for which doubling the computing power has required an increasing amount of work between the 1970s and today. In the current framework, this could be modelled with a lower value of λ resulting in smaller innovation increments. The first row of Figures 9 and 8 presents the effect of a change of λ around the calibrated benchmark.

Households being heterogeneous in their wealth, a positive shock $\Delta\eta$ on the R&D function curvature is interpreted as a "scale-biased" technical change, as it increases the entrepreneurial success rate of wealthy households relative to households who can only afford a lesser entrepreneurial investment. Autor et al. (2020) interpret the fall in labour share observed in recent decades as a consequence of a rise of "superstar firms", a phenomenon that would be attributable to the skewness of the productivity distribution among firms. In the current framework, the distribution of productivity across entrepreneurs is governed by η.

The third driver of inequality I consider is entrepreneurs ability (ϕ_e) to turn effort into new firms. Gao et al. (2013) argue that trade liberalization has introduced fixed costs that acted as larger entry barriers. Moreover, a decline in ϕ_e would align with recent observations of a decrease in the rate of initial public offerings (IPOs) relative to sellouts (Irwin et al., 2019; Bowen et al., 2020), whereby a successful venture may not result in the creation of a new firm, but be absorbed by an existing firm instead. The third row of Figure 9 and 8 present the effects of changes in ϕ_e. Finally, the effect of a wealth tax is also examined.

[7] Here, I define savings as the change in wealth resulting from both households' savings decision and entrepreneurial reward.

Figure 8: *Effect of four drivers of inequality on macroeconomic variables of interest*



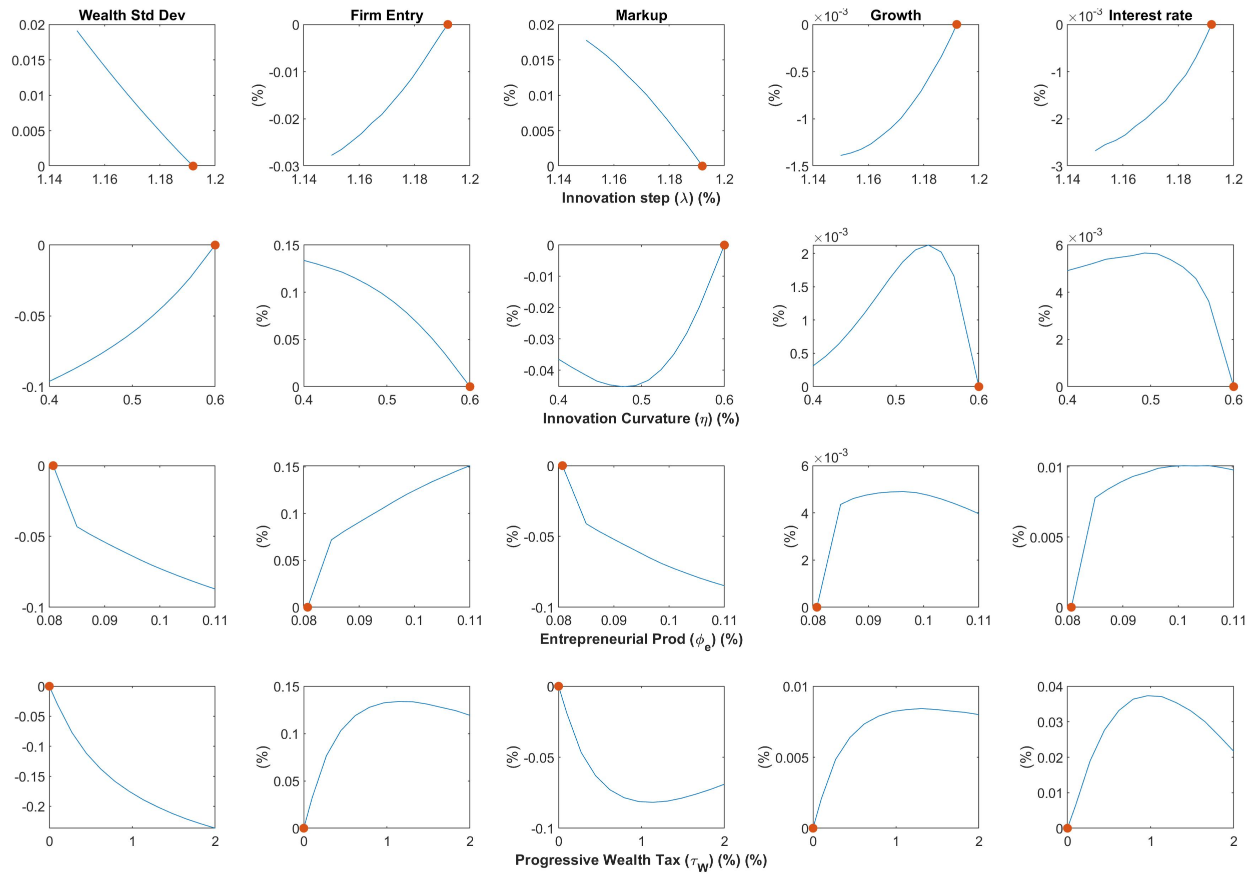
Note: The bold red dot corresponds to the calibrated benchmark.

Household's problem including the wealth tax and the corresponding uniform transfer is presented in (38).

Figure 8 displays the effects of the four inequality drivers on firm entry, firm markup, growth and the interest rate. The solid blue line displays the total effect and the dashed red line presents the effect with the inequality channel shutoff, that is, with the shape of the wealth distribution kept constant. In general, the effect attributed to the change in inequality is small relative to the total effect of its driver, with the exception of the redistributive wealth tax. This should suggest caution in the interpretation of the time series estimation performed in Section 2.

Effects observed in Figure 8 for a positive "scale-biased" technical change ($\Delta\eta$) and a decline in entrepreneurial productivity (ϕ_e) would generate a positive correlation between

Figure 9: *Variations of macroeconomic variables attributed to changes in inequality caused by permanent shocks of four different drivers, locally to the benchmark calibration.*



Note: The bold red dot corresponds to the calibrated benchmark.

changes in wealth inequality and firm markup through the firm entry channel. In both cases, entrepreneurial activity becomes relatively more costly for poorer households, reducing aggregate firm entry and concentrating wealth as a larger share of entrepreneurial success is enjoyed by wealthier households. The reduced entry increases markup and growth decreases as firms competitiveness and innovation decreases.

When innovation step size (λ) declines, however, the ensuing reduction in firm markup for monopolist firms reduces the returns on firm equity, discourages entry and increases the standard deviation of the wealth distribution. While the richest and least risk averse households are disincentivized by the lower payout, the poorest and more risk averse households are encouraged by the lower cost of borrowing. The natural borrowing limit decreases and the variance of the wealth distribution increases. However, the skewness of the distribution also decreases as well as the top wealth shares. The effect of λ on the wealth distribution is more complex than what a single moment can convey, with more heavily indebted households but also a thinner right tail of the distribution.

The effect of a wealth tax is discussed in more detail in the next section. Among the drivers of inequality, it is the only one whose effects on the other variables is in great part attributable to the changes in the shape of the wealth distribution.

Figure 9 displays the part of the effects of the four drivers $(\lambda, \eta, \phi_e, \tau_W)$ attributable to changes in the shape of the wealth distribution. Those effects are calculated as the difference between the total effects and the inequality channel shutoff effects shown respectively as the solid blue line and the dashed red line in Figure 8. Two general observations are in order. First, such inequality channel effects are small. Second, the whole distribution matters. The four exogenous parameters considered alter the wealth distribution in different ways. As a consequence the association between the wealth inequality indicator and macroeconomic variables of interest differs depending on the underlying driver of inequality.

The present analysis will focus on changes happening around the calibrated benchmark highlighted with a bold red dot on both figures 8 and 9. Elasticities of firm entry, markup, growth and the interest rate to changes in the wealth standard deviation around the benchmark are offered in Table 3. An increase in wealth inequality yields a decrease in entrepreneurial activity and an increase in firm markup. The effect is consistent across inequality drivers and with the estimated response functions obtained empirically in Section 2.

Table 3: *Elasticities with respect to changes in the standard deviation of the wealth distribution*

Driver	Firm entry ($\mathcal{Z}$)	Markup ($\mathcal{M}$)	Growth (g)	Interest (r)
Innovation step (λ)	-0.077	0.036	-0.039	-0.024
Innovation curvature (η)	-0.042	0.021	-0.012	-0.009
Entrepreneur Prod. (ϕ_e)	-0.048	0.024	-0.013	-0.010
Wealth Tax (τ_W)	-0.042	0.021	-0.021	-0.009

Note: Elasticities of the variables of interest $(\mathcal{Z}, \mathcal{M}, g, r)$ are evaluated with respect to changes in the indicator of inequality from the calibrated benchmark.

5.3 Policy Implications

The presence of an entrepreneurial entry channel through which wealth inequality affects market power and growth suggests the existence of growth enhancing redistributive fiscal policies.

I consider five policies: three taxes and two subsidies. The three taxes of interest are a corporate profit tax, a tax on capital gains and a progressive wealth tax. For each tax, the government budget is balanced with a uniform transfer to households. I also investigate the effect of two types of subsidies financed with a uniform tax across households: a subsidy on entrepreneurial expenses and a subsidy on firm R&D expenses.

The taxes on capital gains and on wealth affect households decisions differently because, in addition to the returns on savings, an important source of capital gains comes from the creation and sale of a new firm by successful entrepreneurs, whereas the wealth tax is only applied on households savings at any given time. The tax on capital gains thus affects households entrepreneurial efforts in a way that the wealth tax doesn't. The qualitative effect of the wealth tax is unchanged by its progressivity.

The above taxes, transfers and subsidies are incorporated into households and firms problems in (38) to (42). Households problem becomes

$$\psi v(\tilde{a}) = \max_{\tilde{c},z} \left\{ u(\tilde{c}) + v'(\tilde{a})(\dot{\tilde{a}}) + \phi_e l_e^{\eta} \left(v(\tilde{a} + (1-\tau_g)\Pi) - v(\tilde{a}) \right) \right\} \tag{38}$$

s.t.

$$\dot{\tilde{a}} = \tilde{r}\tilde{a} + \tilde{w} - (1-s_e) l_e \tilde{w} + \mathcal{T} - \tilde{c} - \phi z^{\eta} - (\tau_w(\tilde{a}) + \tau_g \tilde{r})\tilde{a}_+ \tag{39}$$

$$\tilde{a} \geq \underline{\mathrm{a}}, \tag{40}$$

where $\tilde{a}_+ \equiv \max\{0, \tilde{a}\}$, τ_g is a tax on capital gains, $\tau_w(\tilde{a})$ is a progressive tax on wealth, s_e is a subsidy on entrepreneurial expenditure, and $\mathcal{T}$ is a uniform transfer (or tax) determined such that the government budget balances. The progressive wealth tax has the form $\tau_w(\tilde{a}) = \left(\frac{\tilde{a}}{a_q}\right)^2 \bar{\tau}_w$ with $\bar{\tau}_w$ being the tax imposed on the maximum wealth bracket and a_q, the wealth percentile threshold for the maximum tax bracket. a_q is set to be the 99.999 percentile.

After incorporation of corporate taxes and subsidies, firms value functions become

$$\tilde{r}\nu_1 = (1-\tau_c)\left(1 - \tfrac{1}{\lambda}\right) + \mathcal{Z}\left(\nu_0 - \nu_1\right), \tag{41}$$

$$\tilde{r}\nu_0 = \max_{l_f} \left\{ -\tilde{w}(1-s_f) l_f + \phi_f l_f^{\eta} \left(\nu_1 - \nu_0\right) - \left(\phi_f l_f^{\eta *} + \tfrac{1}{2}\mathcal{Z}\right) \nu_0 \right\}, \tag{42}$$

where τ_c is a tax on monopolists profits and s_f is a subsidy on firms R&D expenditure.

Table 4 presents the marginal elasticities of wealth standard deviation, firm entry, markup, growth and interest rate associated with each tax and subsidy around zero. All policies considered mitigate wealth inequality in the sense that marginal elasticities for the measure of inequality are negative when tax is null. The progressive wealth tax, and the two subsidies stimulate entrepreneurial activity.

The subsidies stimulate entrepreneurial activity through an incentive effect, while the corporate and the capital gains taxes hinder entrepreneurship. These results regarding subsidies are neither new, nor surprising. Our main result, is that a modest progressive wealth tax also stimulates entrepreneurial activity, and it does so through a wealth redistribution channel. At any given wealth level, households incentive to engage in entrepreneurial activity is reduced by the tax. However, households have a new post-tax wealth level from which, in average, they engage in more entrepreneurial activity than they would at their pre-tax wealth level. This result is new to the literature. It suggests that the redistributive effect of a slight wealth tax could stimulate entrepreneurial activity, increase business competitiveness and stimulate growth despite the distortionary effect of the tax itself.

Guvenen et al. (2019) also find the combination of a positive wealth tax and a negative tax on capital gains to be growth enhancing, in a framework with heterogeneous entrepreneurs differing in ability. Their result stems from the heterogeneity in entrepreneurs ability to turn effort into wealth. The wealth tax enlarges the tax base and the subsidy on capital gains incentivizes the more productive agents to spend greater entrepreneurial efforts. In contrast, in the present framework, all households are ex-ante identical, and they differ only in their wealth level. The entrepreneurship and growth enhancing properties of the wealth tax I observe is solely due to the wealth redistribution channel.

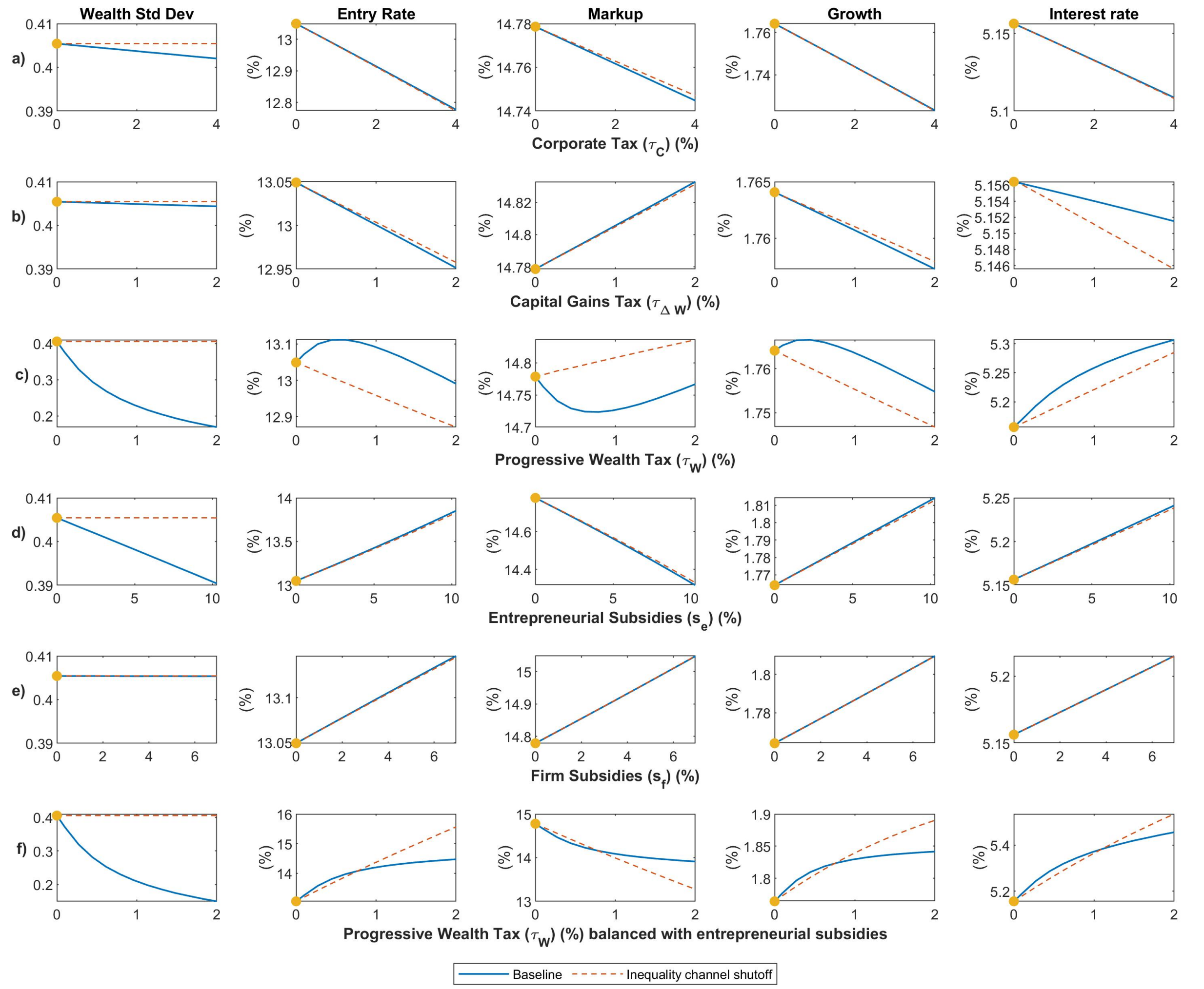
Table 4: *Marginal elasticity of wealth dispersion, firm entry, markup, growth and interest rate to changes in fiscal policies*

	Wealth Std Dev	Firm entry	Markup	Growth	Interest rate
Corporate Tax	-0.598	-1.518	-0.169	-1.687	-0.671
Capital Gains Tax	-0.028	-0.087	0.043	-0.045	-0.011
Progressive Wealth Tax	-89.636	2.070	-1.287	0.799	2.966
Entrepreneurial Subsidy	-10.258	16.438	-8.297	8.110	4.662
Firm Subsidy	-0.092	0.806	1.882	2.687	1.190
Progressive Wealth Tax and Entrepreneurial Subsidy	-102.620	19.031	-9.873	9.021	7.773

Note: Marginal elasticities are computed with respect to taxes as a share of GDP both for tax policies and the subsidies funded by a uniform tax.

Figure 10 presents the effects of all policies examined. The solid blue curve displays the total policy effect, which is decomposed into two parts. The dotted red line displays the incentive effect of the policy, obtained with the wealth distribution channel being shut off by maintaining the shape of the wealth distribution unchanged. The difference between the total policy effect and the incentive effect is the redistributive effect. A more detailed description of the decomposition method accompanies Figure 8 above.

Figure 10: *Effects of policies on wealth inequality, firm entry, markup, growth and interest rate.*



Note: Rows (a) to (e) present the effect of the five policies separately. Row (f) presents the effect of the progressive wealth tax and the entrepreneurial subsidy combined, with the x-axis representing the tax of the highest wealth bracket.

The wealth tax is the only tax, among the considered policies, for which there exists a range in which the positive redistribution effect on entrepreneurial activity overcomes the negative incentive effect of the tax. Households at any given wealth level are disincentivized to invest by the tax. However, the redistribution is such that households have, in average, a wealth level at which they are more inclined to engage in entrepreneurship. Such redistribution effect is also observed with the corporate tax, but it is not strong enough to overcome the corporate tax distortion. The redistribution scheme resulting from the tax on capital gains is small in amplitude, and it further reduces entrepreneurial activity by creating an important distortion around zero wealth and redistributing median household wealth level below that level.

As shown in Table 4 and Figure 10, markups are mitigated by the corporate tax, the progressive wealth tax and the entrepreneurial subsidy through different mechanisms. According to (20), BGP markups depend on two variables: the rate of entry of new firms created by entrepreneurs and the R&D expenditure by toe-to-toe firms in levelled industries. The corporate tax reduces the incentive towards innovation in levelled industries. The decrease in markup is thus accompanied by a decrease in growth. On the other hand, the progressive wealth tax and the entrepreneurial subsidy decrease firm markup by increasing the entry rate of new firms catching up with monopolists. In this case, the decrease in markup is a consequence of an increase in business competitivity and it results in a higher growth rate.

The R&D subsidy of established firms in levelled industries also stimulates growth, however it increases the aggregate firm markup as successful innovative firms gain market power in their respective industries.The effect of the firm subsidy on wealth inequality and markup highlights the difference in the effect of subsidizing an entrepreneur that is set to catch up with the industry leader (entrepreneurial subsidy) versus a firm that is innovating to gain industry leadership (incumbent firm subsidy). Although this distinction may not always be easy to make in practice, the difference in the effect on inequality and business competitiveness is unequivocal, at least under the assumptions underlying the present study, and carries relevant policy implications.

The differences in the effects of entrepreneurial and incumbent firms subsidies provide a theoretical justification for the preferential treatment given to new firms versus incumbent firms in the distribution of subsidies. A large body of literature on the effect of R&D subsidies on private R&D expenditure (reviewed by Zúñiga-Vicente et al. (2014)) shows mixed results and some authors question the relevance of giving a preferential treatment to smaller firms (Holtz-Eakin, 2000) or encouraging individual entrepreneurship altogether (Acs et al., 2016) citing poor theoretical foundation and lack of clear empirical support for such subsidies. If a policy maker is not solely concerned with growth, the present paper provides theoretical justification for a policy maker to consider providing a preferential treatment to entering firms depending on their level of concern with households wealth inequality and business competitiveness.

In this paper, I assume that the entrepreneurial and R&D efforts are public information and that their outcomes follow set innovation functions described in (14) and (21). Entrepreneurship subsidies are not without issues of their own and a detailed examination of their desirability is beyond the scope of this work. Issues include moral hazard (Paulson et al., 2006) and crowding out effects reported in a number of studies reviewed by Zúñiga-Vicente et al. (2014). For that reason, I restricted the analysis to small subsidies with the assumption that some room exists for the policy maker to ease business formation indirectly by, for instance, alleviating some administrative and legal burdens.

In summary, I demonstrate the existence of a redistributive wealth tax that would enhance firm entry, business competitiveness and growth through a wealth redistribution channel. The effect of such a tax on growth is small and caution is required as the negative incentive effect overcomes the redistribution effect as the tax is increased beyond it's optimal level. As a complement to such a tax, it is interesting to note that an entrepreneurial subsidy accomplishes similar results. The combined effect of those two policies is also examined and reported at the bottom of Table 4 and Figure 10, where the wealth tax provides for the entrepreneurial subsidy.

6 Conclusions

In this paper, I investigate how household wealth distribution matters for firms competitiveness. I present time series evidence to the effect that firms' markup responds positively to wealth concentration via the channel of firm entry, and I offer a theoretical framework providing a new perspective on the relationship between household wealth inequality and firms market power. The step-by-step model of technological innovation presented in this paper features entrepreneurial activity as the driver of firm entry in was is otherwise an endogenous growth model of competing firms. The new framework produces both an endogenous distribution of household wealth and a measure of firm markup.

The calibrated framework predicts that a mean-preserving spread in the wealth distribution results in a decrease in entrepreneurial activity and an increase in incumbent firms' aggregate market power. The result is robust to the driver of inequality locally to the calibrated benchmark and is supported by time series evidence. Interestingly, this model generates a wealth distribution with a long right tail without relying on retirement, bequests, precautionary savings or non-homothetic utility function. Risk aversion and the endogenous propensity to engage in risky entrepreneurial activity are the main mechanisms generating the long right tail of the wealth distribution.

In addition to the theoretical framework, I offer evidence from estimation of a structural panel VAR using annual data from 8 OECD countries. I find that firm markup's response to inequality shocks transiting via the channel of firm entry is positive and significant. However, the contribution of inequality shocks to the variance of firm markup is small at around 5%. To conduct this exercise, I used a bootstrap-based bias-corrected maximum likelihood estimation method never applied to dynamic panels before. Moreover, I present a new response decomposition method to quantify the contribution of the firm entry channel to the response of firm markup to an inequality shock.

It is worth noting that both the theoretical model and the time series exercise also produce evidence in support of the hypothesis of inequality responding positively to an increase in market power as suggested by many authors (Comanor and Smiley, 1975; Ennis et al., 2019; Furman and Orszag, 2015; Stiglitz, 2017; Brun and González, 2017).

Finally, I examine a selection of fiscal policies and discuss two main policy implications. First, consistent with Cozzi (2018) and Conesa et al. (2009), I find that under the calibrated benchmark, a slight wealth tax is growth-enhancing due to the uninsurable nature of entrepreneurs risk. I decompose the effect of the tax into an incentive effect that modifies agents decisions conditional on their relative wealth and a redistributive effect that stems from the change in the shape of the wealth distribution. I find that the incentive effect of a small wealth tax on growth is negative, but is overcome by the positive redistributive effect.

An entrepreneurial subsidy also decreases wealth inequality and market power while stimulating firm entry and growth. A R&D subsidy for incumbent firms also stimulates growth at the expense of greater firm market power and wealth concentration. This observation highlights the importance of subsidies targeting and provides a theoretical justification for the preferential treatment often given to new firms in subsidy allocation.

References

Achdou, Y., Han, J., Lasry, J.-M., Lions, P.-L., and Moll, B. (2020). Income and wealth distribution in macroeconomics: A Continuous-Time approach. Working Paper 23732, National Bureau of Economic Research.

Acs, Z., Åstebro, T., Audretsch, D., and Robinson, D. T. (2016). Public policy to promote entrepreneurship: a call to arms. *Small Business Economics*, 47(1):35–51.

Aghion, P., Banerjee, A., and Piketty, T. (1999). Dualism and macroeconomic volatility. *The quarterly journal of economics*, pages 1359–1397.

Aghion, P., Harris, C., Howitt, P., and Vickers, J. (2001). Competition, imitation and growth with Step-by-Step innovation. *The Review of economic studies*, 68(3):467–492.

Akcigit, U. and Ates, S. T. (2019). Ten facts on declining business dynamism and lessons from endogenous growth theory. Working Paper 25755, National Bureau of Economic Research (NBER).

Alvarez, J. and Arellano, M. (2021). Robust likelihood estimation of dynamic panel data models. *Journal of econometrics*.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The quarterly journal of economics*, 135(2):645–709.

Bai, J. (2013). Fixed-effects dynamic panel models, a factor analytical method. *Econometrica*, 81(1):285–314.

Banerjee, A. V. and Newman, A. F. (1994). Poverty, incentives, and development. *The American economic review*, 84(2):211–215.

Bernanke, B. S., Gertler, M., Watson, M., Sims, C. A., and Friedman, B. M. (1997). Systematic monetary policy and the effects of oil price shocks. *Brookings papers on economic activity*, 1997(1):91–157.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *The American economic review*, 110(4):1104–1144.

Bowen, III, D. E., Frésard, L., and Hoberg, G. (2020). Rapidly evolving technologies and startup exits. Technical Report 19-22, Swiss Finance Institute.

Brun, L. and González, I. (2017). Tobin's Q and inequality. Mimeo.

Caggese, A. (2012). Entrepreneurial risk, investment, and innovation. *Journal of financial economics*, 106(2):287–307.

Cavenaile, L., Celik, M. A., and Tian, X. (2020). Are markups too high? competition, strategic innovation, and industry dynamics. Mimeo.

Comanor, W. S. and Smiley, R. H. (1975). Monopoly and the distribution of wealth. *The quarterly journal of economics*, 89(2):177–194.

Conesa, J. C., Kitao, S., and Krueger, D. (2009). Taxing capital? not a bad idea after all! *The American economic review*, 99(1):25–48.

Cozzi, M. (2018). Optimal capital taxation with incomplete markets and schumpeterian growth. Working paper, Department of Economics, University of Victoria.

Davidson, R. and MacKinnon, J. G. (2018). Econometric theory and method, second edition.

Davis, S. J., Haltiwanger, J., Jarmin, R., and Miranda, J. (2007). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. In *NBER Macroeconomics Annual 2006, Volume 21*, pages 107–180. MIT Press.

Davis, S. J. and Haltiwanger, J. C. (2019). Dynamism diminished: The role of housing markets and credit conditions. Working Paper 25466, National Bureau of Economic Research.

De Loecker, J., Eeckhout, J., and Mongey, S. (2021). Quantifying market power and business dynamism in the macroeconomy. Working Paper 28761, National Bureau of Economic Research.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *The quarterly journal of economics*, 135(2):561–644.

Eggertsson, G. B., Lancastre, M., and Summers, L. H. (2019). Aging, output per capita, and secular stagnation. *American Economic Review: Insights*, 1(3):325–342.

Eggertsson, G. B., Robbins, J. A., and Wold, E. G. (2018). Kaldor and piketty's facts: The rise of monopoly power in the united states. Working Paper 24287, National Bureau of Economic Research.

Eichenbaum, M. and Fisher, J. D. M. (2007). Estimating the frequency of price reoptimization in calvo-style models. *Journal of monetary economics*, 54(7):2032–2047.

Ennis, S. F., Gonzaga, P., and Pike, C. (2019). Inequality: A hidden cost of market power. *Oxford Review of Economic Policy*, 35(3):518–549.

Fazzari, S., Hubbard, R. G., and Petersen, B. (1987). Financing constraints and corporate investment. Working Paper 2387, National Bureau of Economic Research.

Foellmi, R. and Zweimüller, J. (2004). Inequality, market power, and product diversity. *Economics letters*, 82(1):139–145.

Foellmi, R. and Zweimüller, J. (2006). Income distribution and Demand-Induced innovations. *The Review of economic studies*, 73(4):941–960.

Furman, J. and Orszag, P. (2015). A firm-level perspective on the role of rents in the rise in inequality. "A Just Society" Centennial Event.

Gandelman, N. and Hernandez-Murillo, R. (2015). Risk aversion at the country level. Mimeo.

Gao, X., Ritter, J. R., and Zhu, Z. (2013). Where have all the IPOs gone? *Journal of Financial and Quantitative Analysis*, 48(6):1663–1692.

García-Peñalosa, C. and Wen, J.-F. (2008). Redistribution and entrepreneurship with schumpeterian growth. *Journal of Economic Growth*, 13(1):57–80.

Gertler, M. and Gilchrist, S. (1994). Monetary policy, business cycles, and the behavior of small manufacturing firms. *The quarterly journal of economics*, 109(2):309–340.

Gilchrist, S. and Himmelberg, C. P. (1995). Evidence on the role of cash flow for investment. *Journal of monetary economics*, 36(3):541–572.

Gutierrez, G. and Philippon, T. (2017). Investmentless growth: An empirical investigation. *Brookings papers on economic activity*, pages 89–170.

Guvenen, F., Kambourov, G., Kuruscu, B., Ocampo-Diaz, S., and Chen, D. (2019). Use it or lose it: Efficiency gains from wealth taxation. Working Paper 26284, National Bureau of Economic Research.

Holtz-Eakin, D. (2000). Public policy toward entrepreneurship. *Small Business Economics*, 15(4):283–291.

Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *Journal of the American Statistical Association*, 75(371):602–608.

Hurwicz, L. (1950). Least squares bias in time series. *Statistical inference in dynamic economic models*, 10:365–383.

Irwin, K. C., Gilstrap, C. M., Drnevich, P. L., and Tudor, C. M. (2019). From start-up to acquisition: Implications of financial investment trends for small- to medium-sized high-tech enterprises. *Journal of Small Business Strategy*, 29(2):22–43.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press.

Kilian, L. and Lewis, L. T. (2011). Does the fed respond to oil price shocks? *The Economic journal*, 121(555):1047–1072.

Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of econometrics*, 68(1):53–78.

Kortum, S. (1993). Equilibrium R&D and the patent-R&D ratio: U.S. evidence. *The American economic review*, 83(2):450–457.

MacKinnon, J. G. and Smith, A. A. (1998). Approximate bias correction in econometrics. *Journal of econometrics*, 85(2):205–230.

Osterwald-Lenum, M. (1992). A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics1. *Oxford bulletin of economics and statistics*, 54(3):461–472.

Paulson, A. L., Townsend, R. M., and Karaivanov, A. (2006). Distinguishing limited liability from moral hazard in a model of entrepreneurship. *The journal of political economy*, 114(1):100–144.

Philippon, T. (2019). *The Great Reversal: How America gave up on free markets*. Harvard University Press.

Phillips, P. C. B. and Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. *Journal of econometrics*, 137(1):162–188.

Quadrini, V. (2000). Entrepreneurship, saving, and social mobility. *Review of economic dynamics*, 3(1):1–40.

Raval, D. (2020). Testing the production approach to markup estimation. Mimeo.

Romano, J. P. and Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282.

Sims, C. A. (2000). Using a likelihood perspective to sharpen econometric discourse: Three examples. *Journal of econometrics*, 95(2):443–462.

Sims, C. A. and Zha, T. (2006). Does monetary policy generate recessions? *Macroeconomic dynamics*, 10(02):231–272.

Stiglitz, J. (2017). Inequality, stagnation, and market power. Working paper, Roosevelt Institute Working Paper.

Traina, J. (2018). Is aggregate market power increasing? production trends using financial statements. Mimeo.

Zúñiga-Vicente, J. Á., Alonso-Borrego, C., Forcadell, F. J., and Galán, J. I. (2014). Assessing the effect of public subsidies on firm R&D investment: A survey. *Journal of economic surveys*, 28(1):36–67.

A Methodological Complement

A.1 Bootstrap-Based Bias Corrected Dynamic Panel Estimator

I estimate (1) by maximum likelihood, conditioning on initial values of the dependent variables and followed the iterative bias correction method suggested by MacKinnon and Smith (1998). I provide the bias-correction algorithm here together with simulations results.

Let B be the set of coefficients to be estimated and $\hat{B}$ be the uncorrected ML estimate. The objective of the bias-correction procedure is to find a corrected estimate $\tilde{B}$ such that if the DGP were to follow $\tilde{B}$, the biased estimate $\hat{B}$ would most likely be obtained. The MacKinnon and Smith (1998)'s procedure follows the following steps:

1. Compute the uncorrected ML estimator $\hat{B}$ and denote it $\hat{B}^{(0)}$
2. For each iteration $i \in \{0, ..., I\}$, generate K bootstrap samples, taking $\hat{B}^{(i)}$ as the true parameters and denote the K estimates from the bootstrap samples as $\hat{B}_j^{(i)}$ ($j \in \{0, ..., K\}$)
3. Define the bootstrap estimate $\bar{B}^{(i)}$ as the median of the $\hat{B}_j^{(i)}$'s
4. Compute the bias $\Delta B^{(i)} = \bar{B}^{(i)} - \hat{B}$
5. Subtract a share $\alpha \in (0, 1)$ of the bias from the biased estimate and obtain the partially corrected estimate $\hat{B}^{(i+1)} = \hat{B}^{(i)} - \alpha \Delta B^{(i)}$
6. Repeat steps 2-5 until the absolute value of all elements of the bias $\Delta B^{(i)}$ are no larger than some tolerance value ϵ.

The convergence speed parameter $\alpha \in (0, 1)$ is chosen to be small enough to allow convergence. Following only steps 1-5 with $\alpha = 1$ yields what is referred below as the one-step bias-correction procedure.

In the specific context of this estimation, the MacKinnon and Smith (1998) procedure works remarkably well despite the high dimensionality of the parameter space. Table 5 presents a series of bias-correction performance measures obtained from 1000 Monte Carlo repetitions taking the bias-corrected estimate as the true DGP and using resampled residuals as simulated errors. For each thus generated sample, three estimates are computed according to different bias-correction procedures: No correction, one-step correction and iterated correction. Performances are assessed by comparing the mean absolute bias (MAB), the root mean square bias (RMSB) and the estimates' average standard deviation (ASD) across Monte Carlo samples.

Since estimates are biased towards stationarity, it is interesting to examine separately the bias-correction performance related to the first order persistence terms on the diagonal of β_1, the first matrix of AR coefficients. One can observe that the bias is more important on the diagonal of the first order coefficients and that the two bias-correction methods realize important gains. Notably, the bias on the first diagonal is corrected from the RMSB being twice the ASD when left uncorrected to being less than half the ASD with the iterated correction. Relative to the true values, the bias is reduced from 27.8% to 34.8% of the true parameter along the first diagonal when uncorrected to 3.0% to 10.6% when the iterated correction is applied. Moreover, the gain in terms of bias comes at a very low cost in terms of additional variance.

Table 5: *Bias-correction performance measures*

	No Correction	1-step correction	Iterated correction
MAB	0.0819	0.0406	0.0302
RMSB	0.1578	0.0833	0.0655
RMSB (1st diagonal)	0.2261	0.0887	0.0535
ASD	0.3022	0.3007	0.3053
ASD (1st diagonal)	0.1114	0.1191	0.1240

Note: Measures are applied to the bias-correction of the autoregressive coefficients (B). The mean absolute bias (MAB), root mean square bias (RMSB) and the estimates' average standard deviation (ASD) are obtained from 1000 Monte Carlo repetitions. Measures of performance indicators applied only to diagonal elements of the first AR coefficients matrix β_1 are also provided.

A.2 Lag Selection

Lag selection is determined by minimization of the Bayesian-Schwarz information criterion (BIC) conditional on the absence of significant serial correlation among residuals. Table 7 provides the AIC, BIC and HQC information criteria for a selection of models differing by their lags set $\mathcal{L}$ and including lags up to 5 years, a common duration for mortgage contracts that may be a significant part of small business owners credit. For each country, a portmanteau test for residual autocorrelation is performed and I reject the null of no autocorrelation at a conservative level of 0.2. The test statistic (Hosking, 1980) is given by

$$Q_h = T^2 \sum_{j=1}^{h} \frac{1}{T-j} tr(\hat{C}_j' \hat{C}_0^{-1} \hat{C}_j \hat{C}_0^{-1}), \tag{43}$$

where $\hat{C}_j = \frac{1}{T} \sum_{t=j+1}^{T} \hat{\epsilon}_t \hat{\epsilon_{t-j}'}$ and $\hat{\epsilon}_t$ are the residuals from the estimation of (1). P-values are obtained using a bootstrap procedure described below.

The test is applied to each country separately and a unique p-value is obtained following a bootstrap multiple testing adjustment procedure proposed in Davidson and MacKinnon (2018) that is similar to what Romano and Wolf (2005) describe as the bootstrap-based single-step method for multiple testing. Bootstrap samples are constructed under the null of no residual autocorrelation by resampling residuals within countries and thereby breaking any autocorrelation patterns. For each sample, the test is performed separately for each country and the minimum p-value from those tests is considered. The p-value adjusted for multiple testing is given by the rate at which the bootstrap minimum p-values are lower than the original test p-value. The adjusted p-value is thus computed for each autocorrelation lag considered and displayed in columns titled L1 to L10 of Table 6. The procedure is repeated across both dimensions (the eight countries and the ten autocorrelation lags), and the unique adjusted p-value is presented in the last column of Table 6.

The most parsimonious model minimizing the BIC criterion while presenting a p-value over 0.15 for the autocorrelation test has lags set $\mathcal{L} = \{1, 5\}$.

Table 6: *P-values for the porte-manteau test of no autocorrelation*

$\mathcal{L}$	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	p-value
1	0.2064	0.0822	0.0581	0.0080	0.0160	0.0060	0.0180	0.0120	0.0100	0.0140	0.0301
12	0.4790	0.6212	0.2505	0.1283	0.2846	0.1202	0.0561	0.0521	0.0581	0.0681	0.1002
13	0.6934	0.5110	0.1423	0.0641	0.0521	0.0020	0.0281	0.0100	0.0160	0.0601	0.0180
14	0.0521	0.0521	0.0601	0.1242	0.2465	0.2345	0.0982	0.0741	0.0922	0.0641	0.1343
15	0.0822	0.0321	0.0721	0.0621	0.2004	0.2445	0.2305	0.2144	0.2745	0.4549	0.2124
123	0.8597	0.2204	0.0782	0.1363	0.0982	0.0381	0.0461	0.0261	0.0301	0.0301	0.0521
124	0.8557	0.1884	0.1303	0.2545	0.7575	0.7174	0.2685	0.1984	0.2184	0.3908	0.4128
125	0.2445	0.0261	0.1002	0.0721	0.2204	0.3507	0.4950	0.5932	0.7315	0.9699	0.1884
134	0.1122	0.0681	0.0741	0.2285	0.5832	0.4609	0.4729	0.2184	0.2766	0.2585	0.3126
135	0.3026	0.0681	0.1523	0.1022	0.2926	0.1142	0.1042	0.1363	0.2204	0.4148	0.2465

Note: Columns L1 to L10 present the p-values for the test performed with autocorrelation lags 1 to 10 respectively, adjusted for multiple testing across countries. The last columns presents the p-value adjusted for multiple testing across countries and autocorrelation lags.

Table 7: *Information criteria for a selection $\mathcal{L}$ of models*

$\mathcal{L}$	AIC	BIC	HQC
1	12.50	16.42	13.10
12	12.76	17.65	13.51
13	12.83	17.92	13.61
14	12.92	18.21	13.73
15	13.01	18.54	13.85
123	13.04	18.97	13.95
124	13.15	19.32	14.08
125	13.25	19.70	14.23
134	13.14	19.32	14.08
135	13.25	19.69	14.22

A.3 Cointegration Rank

I perform a simple multivariate test for the number of cointegration relations on the panel vector error correction (PVEC) model representation of (1). The PVEC model with unrestricted constants and trends is given by

$$\Delta y_{it} = \eta_i + \delta_i t + \Pi y_{i,t-1} + \gamma_1 \Delta y_{i,t-1} + ... + \gamma_{p-1} \Delta y_{i,t-p+1} + \epsilon_{it}, \quad t \in (t_{0i}, ..., T_i), \tag{44}$$

where $\Pi = (\beta_1 + ... + \beta_p - I)$ and $\gamma_i = \sum_{j=i+1}^{p} \beta_j$.

Johansen (1995)'s trace and maximum tests are performed from the eigenvalues associated with 44 to test for the cointegration rank (r) of the matrix Π. The pseudo-likelihood ratio statistics are denoted $\mathrm{LR}_{\mathrm{trace}}$ and $\mathrm{LR}_{\max}$, and their corresponding critical values for a 10% level test are $\mathrm{CV}_{\mathrm{trace}}$ and $\mathrm{CV}_{\max}$. Results from the two cointegration tests are displayed in Table 8.

Table 8: *Johansen cointegration tests*

Eigenvalues	H_0	$\mathrm{LR}_{\mathrm{trace}}$	$\mathrm{CV}_{\mathrm{trace}}$	$\mathrm{LR}_{\max}$	$\mathrm{CV}_{\max}$
0.3546	rank=0	171.53	43.95	59.56	24.73
0.3169	rank=1	111.97	26.79	51.83	18.60
0.2435	rank=2	60.15	13.33	37.96	12.07
0.1505	rank=3	22.19	2.69	22.19	2.69

Note: Critical values are obtained from Osterwald-Lenum (1992) for 10% level tests.

The trace test rejects the null hypotheses $H_0 : rank(\Pi) = r_0$ against the alternative $H_1 : rank(\Pi) > r_0$ for $r_0 = \{0, 1, 2, 3\}$. Likewise, the maximum test rejects the null hypotheses $H_0 : rank(\Pi) = r_0$ against the alternative $H_1 : rank(\Pi) = r_0 + 1$ for $r_0 = \{0, 1, 2, 3\}$.

I conclude that Π has full rank and that the system is thus stationary and I estimate a PVAR in level. Since all variables are rates, this result is in line with economic theory, at least in the long-run.

B Derivations

B.1 Deriving Entrepreneurs' Problem

Households maximize their lifetime expected utility by choice of a consumption and an R&D intensity paths subject to a budget constraint and a borrowing constraint. Their optimization problem is given by

$$V_t(a_t) = \max_{\{c_t, l_t^e\}} E_t \left[\int_t^T e^{-\rho(\tau-t)} u(c_\tau) d\tau + e^{-\rho T} V_T(a_T + \Pi_T) \right] \tag{45}$$

s.t.

$$\dot{a}_t = r_t a_t + w_t(1 - l_t^e) - c_t \tag{46}$$

$$a_t \geq \underline{\mathrm{a}} Y_t. \tag{47}$$

Household wealth at time t is denoted a_t and the expectation is taken over the uncertain project completion time T, which depends on the research effort path $\{l_t^e\}$. The borrowing constraint evolves in time and minus $\underline{\mathrm{a}}$ is the maximum share of Y_t one is allowed to borrow.

We assume that R& D is a Poisson process and that at each instant the density probability of success is given by $z_t(l_t^e)e^{-z_t(l_t^e)t}$ where $z_t(l_t^e) = \phi_e(l_t^e)^\eta$. To simplify notation, the density probability of success is simply denoted by z_t below. The value function at time t can then be written as

$$V_t(a_t) = \max_{\{c_t, l_t^e\}} \int_t^\infty e^{-\rho(\tau-t)} \left[\Pr(T \geq \tau) \cdot u(c_\tau) + \Pr(T = \tau) \cdot V_\tau(a_\tau + \Pi_\tau) \right] d\tau, \tag{48}$$

with the probability density function given by $\Pr(T = \tau) = z_\tau e^{-z_\tau(\tau-t)}$ and the associated cumulative density given by $\Pr(T \geq \tau) = 1 - \int_t^\tau z_s e^{-z_s(s-t)} ds$. When z_τ is constant over some interval, the above pdf and cdf conditional on $T > t$ are given by

$$\Pr(T = \tau) = z_\tau e^{-z_\tau(\tau-t)}$$

$$\Pr(T \geq \tau) = e^{-z_\tau(\tau-t)}.$$

B.1.1 Derivation of the Hamilton-Jacobi-Bellman (HJB) Equation

Consider $\mathcal{T} \equiv t + \Delta t$ and let Δt be a time interval small enough that z_t is constant over that interval. The integral can be evaluated separately before and after time $\mathcal{T}$ and the value function becomes

$$\begin{aligned} V_t(a_t) = \max_{\{c_t, l_t^e\}} & \int_t^{\mathcal{T}} e^{-(\rho+z_\tau)(\tau-t)} \left[u(c_\tau) + z_\tau V_\tau(a_\tau + \Pi_\tau)\right] d\tau \\ & + \int_{\mathcal{T}}^{\infty} e^{-(\rho+z_{\mathcal{T}})(\mathcal{T}-t)} e^{-\rho(\tau-\mathcal{T})} \left[\Pr(T \geq \tau) \cdot u(c_\tau) + \Pr(T = \tau) \cdot V_\tau(a_\tau + \Pi_\tau)\right] d\tau, \end{aligned}$$

using the fact that z_t is constant between t and $\mathcal{T}$ and thus, $\Pr(T \geq \mathcal{T}) = e^{-z_{\mathcal{T}}(\mathcal{T}-t)}$. In the second integral, the term $e^{-(\rho+z_{\mathcal{T}})(\mathcal{T}-t)}$ is independent of the integrating variable and can be factored out of the integral, leaving the integral to be nothing else than $V_{\mathcal{T}}(a_{\mathcal{T}})$.

$$V_t(a_t) = \max_{\{c_t, l_t^e\}} \int_t^{\mathcal{T}} e^{-(\rho+z_\tau)(\tau-t)} \left[u(c_\tau) + z_\tau V_\tau(a_\tau + \Pi_\tau)\right] d\tau + e^{-(\rho+z_{\mathcal{T}})(\mathcal{T}-t)} V_{\mathcal{T}}(a_{\mathcal{T}}).$$

Taking the derivative with respect to $\mathcal{T}$ and using an envelop theorem argument to ignore the derivative of $z_{\mathcal{T}}$ yields[8]

$$\begin{aligned} 0 =& e^{-(z_{\mathcal{T}}+\rho)(\mathcal{T}-t)} \left[u(c_{\mathcal{T}}^*) + z_{\mathcal{T}} V_{\mathcal{T}}(a_{\mathcal{T}} + \Pi_{\mathcal{T}})\right] - (\rho + z_{\mathcal{T}}^*) e^{-(\rho+z_{\mathcal{T}}^*)(\mathcal{T}-t)} \\ & + e^{-(\rho+z_{\mathcal{T}}^*)(\mathcal{T}-t)} \left(V_{\mathcal{T}}'(a_{\mathcal{T}})\dot{a}_{\mathcal{T}} + \dot{V}_{\mathcal{T}}(a_{\mathcal{T}})\right) \\ 0 =& u(c_{\mathcal{T}}^*) + z_{\mathcal{T}}^* V_{\mathcal{T}}(a_{\mathcal{T}} + \Pi_{\mathcal{T}}) - (\rho + z_{\mathcal{T}}^*) V_{\mathcal{T}}(a_{\mathcal{T}}) + V_{\mathcal{T}}'(a_{\mathcal{T}})\dot{a}_{\mathcal{T}} + \dot{V}_{\mathcal{T}}(a_{\mathcal{T}}), \end{aligned}$$

where $\{c_t^*\}$ and $\{z_t^*\}$ are maximizing paths for consumption and research intensity. Since t is arbitrary, so is $\mathcal{T}$ and the above can be generalized to any time t. Rearranging the terms and substituting the law of motion for wealth yields the HJB equation.

$$\rho V_t(a_t) - \dot{V}_t(a_t) = u(c_t^*) + V_t'(a_t)\dot{a}_t + z_t^* \left[V_t(a_t + \Pi_t) - V_t(a_t)\right],$$

[8] An argument must be made about the preservation of the max.

subject to (46) and (47). Since $\{c_t^*\}$ and $\{z_t^*\}$ maximize $V_t(a_t)$, the HJB equation can be written

$$\rho V_t(a_t) - \dot{V}_t(a_t) = \max_{\{c_t, z_t\}} u(c_t) + V_t'(a_t)\dot{a}_t + z_t \left[V_t(a_t + \Pi_t) - V_t(a_t) \right]. \tag{49}$$

Equation 49 can also be derived from the discrete-time problem of the entrepreneur.

B.1.2 Deriving the HJB Equation from the Discrete-time Problem

Consider a small period length Δ and a discount factor $\beta(\Delta) = e^{-\rho\Delta}$. The probability that an entrepreneur's R&D project completes successfully between times t and $t + \Delta$ is given by $p_t(\Delta) = \int_t^{t+\Delta} z_s e^{-z_s(s-t)} ds$. The discrete-time formulation of the entrepreneurs' problem is given by

$$V_t(a_t) = \max_{c_t, l_t^e} \Delta u(c_t) + \beta(\Delta) \left[p_t(\Delta) V_{t+\Delta}(a_{t+\Delta} + \Pi_{t+\Delta}) + (1 - p_t(\Delta)) V_{t+\Delta}(a_{t+\Delta}) \right] \tag{50}$$

s.t.

$$a_{t+\Delta} = (1 + r_t) a_t + w_t (1 - l_t^e) - c_t \tag{51}$$

$$a_{t+\Delta} \geq \underline{\mathrm{a}} Y_t. \tag{52}$$

The period Δ can be chosen to be arbitrarily small so that the following approximations hold.

$$\beta(\Delta) = e^{-\rho\Delta} \approx 1 - \rho\Delta, \quad p_t \approx 1 - e^{-z_t\Delta} \approx z_t\Delta.$$

Substituting into (50) yields

$$V_t(a_t) = \max_{c_t, z_t} \Delta u(c_t) + (1 - \rho\Delta) \left[z_t \Delta V_{t+\Delta}(a_{t+\Delta} + \Pi_{t+\Delta}) + (1 - z_t\Delta) V_{t+\Delta}(a_{t+\Delta}) \right]$$

subject to (51) and (52). Subtracting $(1-\rho\Delta)V_t(a_t)$ from both sides, we have

$$\rho\Delta V_t(a_t) = \max_{c_t,z_t} \Delta u(c_t) + (1-\rho\Delta)\left[\left(V_{t+\Delta}(a_{t+\Delta}) - V_t(a_t)\right) + z_t\Delta\left(V_{t+\Delta}(a_{t+\Delta}+\Pi_{t+\Delta}) - V_{t+\Delta}(a_{t+\Delta})\right)\right]$$

subject to (51) and (52). Adding and subtracting $V_{t+\Delta}(a_t)$ on the RHS and dividing by Δ, we have

$$\begin{aligned}\rho V_t(a_t) =& \max_{c_t,z_t} u(c_t) + (1-\rho\Delta)\Big[z_t\left(V_{t+\Delta}(a_{t+\Delta}+\Pi_{t+\Delta}) - V_{t+\Delta}(a_{t+\Delta})\right) \\ &+ \frac{V_{t+\Delta}(a_{t+\Delta}) - V_{t+\Delta}(a_t)}{\Delta} + \frac{V_{t+\Delta}(a_t) - V_t(a_t)}{\Delta}\Big]\end{aligned}$$

subject to (51) and (52). Taking the limit as $\Delta \to 0$, yields

$$\rho V_t(a_t) = \max_{c_t,z_t} u(c_t) + \frac{\partial V_t(a_t)}{\partial a_t}\dot{a}_t + \frac{\partial V_t(a_t)}{\partial t} + \phi_e(l_t^e)^\eta\left(V_t(a_t+\Pi_t) - V_t(a_t)\right), \tag{53}$$

subject to (51) and (52).

B.1.3 A Time-Independent Formulation for the HJB Equation

The value function V_t defined above depends on time only through the control and state variables and through the profit Π_t which is given by the value of the toe-to-toe firm ($\mathcal{V}_{0,t}$). On a BGP, the interest rate and the growth rate are constant. This will allow for a time-independent formulation of the HJB equation.

We define the normalized variables $\tilde{c}_t \equiv \frac{c_t}{Y_t}$, $\tilde{a}_t \equiv \frac{a_t}{Y_t}$, $\tilde{w}_t \equiv \frac{w_t}{Y_t}$, $\tilde{\Pi} \equiv \frac{\mathcal{V}_{0,t}}{Y_t} = \nu_0$, and let $\tilde{V}_t(\tilde{a}_t) \equiv V_t(a_t)$ take the normalized wealth ($\tilde{a}_t$) as argument. Then,

- $$\begin{aligned}\frac{\partial V_t(a_t)}{\partial t} =& \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial t} + \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial \tilde{a}_t}\frac{\partial}{\partial t}\left(\frac{1}{Y_t}\right)a_t \\ =& \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial t} - \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial \tilde{a}_t}g\tilde{a}_t;\end{aligned}$$
- $$\frac{\partial V_t(a_t)}{\partial a_t} = \frac{1}{Y_t}\frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial \tilde{a}_t}.$$

Substituting the above derivatives and expressing the HJB equation in terms of the value function $\tilde{V}_t(\tilde{a}_t)$ taking the normalized argument yields

$$\rho \tilde{V}_t(\tilde{a}_t) = \max_{c_t, l_t^e} u(c_t) + \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial \tilde{a}_t} \left(\tilde{r}\tilde{a}_t + \tilde{w}(1 - l_t^e) - \tfrac{c_t}{Y_t} \right) + \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial t} + \phi_e (l_t^e)^\eta \left(\tilde{V}_t(\tilde{a}_t + \tilde{\Pi}) - \tilde{V}_t(\tilde{a}_t) \right), \tag{54}$$

with $\tilde{r} \equiv r - g$. Note that profit $\tilde{\Pi}$ is constant when expressed relative to Y_t. It would be convenient to express consumption in relative terms as well and the following will help reformulate the value function for that purpose.

Suppose $\gamma \neq 1$ then the utility function taking the relative consumption as argument is

$$u(\tilde{c}_t) = \frac{\left(\frac{c_t}{Y_t}\right)^{1-\gamma} - 1}{1-\gamma} = \left(\frac{1}{Y_t}\right)^{1-\gamma} \left(u(c_t) - u(Y_t) \right). \tag{55}$$

It can easily be shown that (55) also holds when $\gamma = 1$. The HJB equation can be formulated in terms of relative consumption by defining $v_t(\cdot)$ such that $\rho v_t(\tilde{a}_t) = \left(\frac{1}{Y_t}\right)^{1-\gamma} \left(\rho \tilde{V}_t(\tilde{a}_t) - u(Y_t) \right) - \frac{g}{\psi}$, with $\psi \equiv \rho - (1-\gamma)g$. Then, the following holds

- $\tilde{V}_t(\tilde{a}_t) = \left(v_t(\tilde{a}_t) + \frac{g}{\rho\psi} \right) Y_t^{1-\gamma} + \frac{1}{\rho} u(Y_t);$
- $$\begin{aligned} \frac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial t} &= Y_t^{1-\gamma} \frac{\partial v_t(\tilde{a}_t)}{\partial t} + \left(v_t(\tilde{a}_t) + \tfrac{g}{\rho\psi} \right) (1-\gamma) Y_t^{-\gamma} \dot{Y}_t + \tfrac{1}{\rho} Y_t^{-\gamma} \dot{Y}_t \\ &= Y_t^{1-\gamma} \left[\frac{\partial v_t(\tilde{a}_t)}{\partial t} + \left(v_t(\tilde{a}_t) + \tfrac{g}{\rho\psi} \right) (1-\gamma) g + \tfrac{g}{\rho} \right]; \end{aligned}$$
- $\dfrac{\partial \tilde{V}_t(\tilde{a}_t)}{\partial \tilde{a}_t} = Y_t^{1-\gamma} \dfrac{\partial v_t(\tilde{a}_t)}{\partial \tilde{a}_t}.$

Substituting the above into (54) yields

$$\begin{aligned} \rho v_t(\tilde{a}_t) &= \max_{\tilde{c}_t, l_t^e} u(\tilde{c}_t) - \tfrac{g}{\psi} + \frac{\partial v_t(\tilde{a}_t)}{\partial \tilde{a}_t} \left(\tilde{r}\tilde{a}_t + \tilde{w}(1 - l_t^e) - \tilde{c}_t \right) \\ &\quad + \left[\frac{\partial v_t(\tilde{a}_t)}{\partial t} + \left(v_t(\tilde{a}_t) + \tfrac{g}{\rho\psi} \right) (1-\gamma) g + \tfrac{g}{\rho} \right] + \phi_e (l_t^e)^\eta \left(v_t(\tilde{a}_t + \tilde{\Pi}) - v_t(\tilde{a}_t) \right) \\ \psi v_t(\tilde{a}_t) &= \max_{\tilde{c}_t, l_t^e} u(\tilde{c}_t) + \frac{\partial v_t(\tilde{a}_t)}{\partial \tilde{a}_t} \left(\tilde{r}\tilde{a}_t + \tilde{w}(1 - l_t^e) - \tilde{c}_t \right) \end{aligned}$$

$$+\frac{\partial v_t(\tilde{a}_t)}{\partial t}+\phi_e(l_t^e)^\eta\left(v_t(\tilde{a}_t+\tilde{\Pi})-v_t(\tilde{a}_t)\right)-\tfrac{g}{\psi}+\tfrac{g}{\rho\psi}(1-\gamma)g+\tfrac{g}{\rho}.$$

The constant terms cancel out and we are left with

$$\begin{aligned}\psi v_t(\tilde{a}_t)=&\max_{\tilde{c}_t,l_t^e}u(\tilde{c}_t)+\frac{\partial v_t(\tilde{a}_t)}{\partial \tilde{a}_t}\left(\tilde{r}\tilde{a}_t+\tilde{w}(1-l_t^e)-\tilde{c}_t\right)\\&+\frac{\partial v_t(\tilde{a}_t)}{\partial t}+\phi_e(l_t^e)^\eta\left(v_t(\tilde{a}_t+\tilde{\Pi})-v_t(\tilde{a}_t)\right).\end{aligned}$$

Note that the normalized value function $v_t(\cdot)$ only depends on time t through the state and control variables. Thus, the function itself is autonomous, the partial derivative with respect to time is zero and the time subscript can be dropped from the value function. For notational convenience, we will also drop the time subscript from the variables where it does not interfere with clarity. The autonomous HJB equation is thus

$$\psi v(\tilde{a})=\max_{\tilde{c},l^e}u(\tilde{c})+\frac{\partial v(\tilde{a})}{\partial \tilde{a}}\left(\tilde{r}\tilde{a}+\tilde{w}(1-l^e)-\tilde{c}\right)+\phi_e(l^e)^\eta\left(v(\tilde{a}+\tilde{\Pi})-v(\tilde{a})\right)\tag{56}$$

subject to

$$\dot{\tilde{a}}=\tilde{r}\tilde{a}+\tilde{w}(1-l^e)-\tilde{c}\tag{57}$$

$$\tilde{a}\geq\underline{\mathrm{a}},\tag{58}$$

where $\tilde{r}\equiv r-g$ is the interest rate applied to the relative wealth.[9]

[9]Note that the effective interest rate applied to $\tilde{a}_t$ is $r-g$ because the relative wealth $\tilde{a}_t$ decreases as Y_t grows. More specifically,

$$\dot{\tilde{a}}_t\equiv\frac{d}{dt}\left(\frac{a_t}{Y_t}\right)=\frac{\dot{a}_t}{Y_t}-\frac{a_t\dot{Y}_t}{Y_t^2}=\left(r\tilde{a}_t+\tilde{w}(1-l^e)-\tilde{c}_t\right)-g\tilde{a}_t.$$

B.2 Aggregate Savings and Wealth Distribution

Relative household wealth $\tilde{a}_t$ evolves according to the savings function $s(\tilde{a}_t, t)$ as

$$d\tilde{a}_t = s_t(\tilde{a}_t)dt. \tag{59}$$

The state of the economy is determined by quality levels $\{k_j\}$ and the relative wealth distribution $f_t(\tilde{a})$. We assume that an entrepreneur's project is not in direct competition with his previous innovation and randomly reassign successful entrepreneurs to a new line j. It follows from this assumption that the distribution of wealth is the same across production lines j and that only the aggregate wealth distribution matters.[10]

Consider the cumulative distribution function

$$\begin{aligned} F_t(\tilde{a}) =& \Pr(\tilde{a}_t \leq \tilde{a}|t) \\ &\text{with} \\ F_t(\underline{\text{a}}) =& 0 \quad \forall t, \\ \lim_{\tilde{a}\to\infty} F_t(\tilde{a}) =& 1 \quad \forall t, \\ f_t(\tilde{a}) \equiv& F_t'(\tilde{a}). \end{aligned}$$

We now proceed to derive a law of motion for the cumulative distribution $F_t(\tilde{a})$. Consider a discrete-time decision process with small period length Δ. Entrepreneurs make their savings decisions according to

$$\tilde{a}_t = \tilde{a}_{t+\Delta} - \Delta s_t(\tilde{a}_{t+\Delta}). \tag{60}$$

Given the wealth distribution at time t, the proportion of households with wealth below $\tilde{a}$ at time $t+\Delta$ is obtained by adding the proportion of entrepreneurs who will be unsuccessful during the time interval Δ and whose wealth is below $\tilde{a} - \Delta\tilde{s}_t(\tilde{a})$ to the proportion of

[10] An alternative assumption would be that entrepreneurs don't know the distribution of wealth in their production line, but that they know the total wealth distribution.

entrepreneurs who will be successful but whose wealth at time t is such that the prize $\tilde{\Pi}$ will be insufficient to increase their wealth level above $\tilde{a}$. Formally,

$$
\begin{aligned}
F_{t+\Delta}(\tilde{a}) =& \int_{\underline{a}}^{\tilde{a}-\Delta s_t(\tilde{a})} (1-\Delta z_t(\zeta)) f_t(\zeta) d\zeta + \int_{\underline{a}}^{\tilde{a}-\Delta s_t(\tilde{a})-\tilde{\Pi}} \Delta z_t(\zeta) f_t(\zeta) d\zeta \\
=& F_t(\tilde{a}-\Delta \tilde{s}_t(\tilde{a})) - \int_{\tilde{a}-\Delta s_t(\tilde{a})-\tilde{\Pi}}^{\tilde{a}-\Delta \tilde{s}_t(\tilde{a})} \Delta z_t(\zeta) f_t(\zeta) d\zeta.
\end{aligned}
$$

Subtracting $F_t(\tilde{a})$ from both sides and dividing by Δ yields

$$
\frac{F_{t+\Delta}(\tilde{a}) - F_t(\tilde{a})}{\Delta} = \frac{F_t(\tilde{a}-\Delta s_t(\tilde{a})) - F_t(\tilde{a})}{\Delta} - \int_{\tilde{a}-\Delta s_t(\tilde{a})-\tilde{\Pi}}^{\tilde{a}-\Delta s_t(\tilde{a})} z_t(\zeta) f_t(\zeta) d\zeta.
$$

Taking the limit as $\Delta \to 0$,

$$
\frac{\partial}{\partial t} F_t(\tilde{a}) = - s_t(\tilde{a}) \frac{\partial}{\partial \tilde{a}} F_t(\tilde{a}) - \int_{\tilde{a}-\tilde{\Pi}}^{\tilde{a}} z_t(\zeta) f_t(\zeta) d\zeta.
$$

Taking the derivative with respect to $\tilde{a}$, we have

$$
\dot{f}_t(\tilde{a}) = -\frac{\partial}{\partial \tilde{a}} \big[s_t(\tilde{a}) f_t(\tilde{a}) \big] - z_t(\tilde{a}) f_t(\tilde{a}) + z_t(\tilde{a}-\tilde{\Pi}) f_t(\tilde{a}-\tilde{\Pi}). \tag{61}
$$

Recognizing that $s_t(\tilde{a}_t) = \dot{\tilde{a}}$,

$$
\dot{f}_t(\tilde{a}) = -\frac{\partial}{\partial \tilde{a}} \big[(\tilde{r}\tilde{a} + \tilde{w} - \tilde{c}(\tilde{a}) - \phi z(\tilde{a})^{\eta}) f_t(\tilde{a}) \big] - z(\tilde{a}) f_t(\tilde{a}) + z(\tilde{a}-\tilde{\Pi}) f_t(\tilde{a}-\tilde{\Pi}). \tag{62}
$$

Since no variables and parameters from the equation above depend on t on a BGP, the solution for $f_t(\cdot)$ is then the same for all t. The t subscripts can be dropped without loss of generality. Using the long notation $z(\tilde{a}) = \phi_e l^e(\tilde{a})^{\eta}$, the time-independent Kolmogorov forward equation for a BGP stationary wealth distribution is given by

$$
0 = -\frac{\partial}{\partial \tilde{a}} \big[(\tilde{r}\tilde{a} + \tilde{w}(1-l^e(\tilde{a})) - \tilde{c}(\tilde{a})) f(\tilde{a}) \big] - \phi_e l^e(\tilde{a})^{\eta} f(\tilde{a}) + \phi_e l^e(\tilde{a})^{\eta} (\tilde{a}-\tilde{\Pi}) f(\tilde{a}-\tilde{\Pi}). \tag{63}
$$

C Proofs

The following lemma will prove useful to prove Proposition 1.

Lemma 1 *Let $\phi_{ij}(L)$ be a power series of lag operators and $\hat{X}_{j,t+k}$, a function of η_{1t}. Let $\Phi_{j1,k}$ be the response of $\hat{X}_j$ to a shock η_{1t} at horizon k. Then,*

$$\frac{d\ \phi_{ij}(L)\hat{X}_{j,t+k}}{d\eta_{1t}} = \phi_{ij}(L)\frac{d\hat{X}_{j,t+k}}{d\eta_{1t}} = \phi_{ij}(L)\Phi_{j1,k}$$

where the lag operator is applied on the horizon k on the RHS.

Proof. Expand the $\phi_{ij}(L)$ power series, and apply the lag operator on the horizon k.

$$\begin{aligned}\frac{\partial\phi_{ij}(L)\hat{X}_{j,t+k}}{\partial\eta_{1t}} &= \phi_{ij,0}\frac{\partial\hat{X}_{j,t+k}}{\partial\eta_{1t}} + \phi_{ij,1}\frac{\partial\hat{X}_{j,t+k-1}}{\partial\eta_{1t}} + ... + \phi_{ij,k-1}\frac{\partial\hat{X}_{j,1-p}}{\partial\eta_{1t}} \\ &= \phi_{ij,0}\Phi_{j1,k} + \phi_{ij,1}\Phi_{j1,k-1} + ... + \phi_{ij,k-1}\Phi_{j1,1} = \phi_{qi}(L)\Phi_{j1,k}\end{aligned}$$

which completes the proof. ■

Proposition 1 is restated here for convenience. The intuition behind the proposition is that the persistence $\Phi_{jj}(L)$ is taken into account in both responses $\Phi_{qj}(L)$ and $\Phi_{ji}(L)$ and is thus computed twice when the two IRF power series are multiplied. A formal proof is offered below.

Proposition 1 (IRF Channel Contribution) *Consider a $n \times 1$ stationary variable $X_t = (X_{qt}, X_{jt}, X_{it}, ...)'$. Let $\Phi_{qi}(L)$ be the power series describing the response of variable X_{qt} to a shock on variable X_{it}. Let $\bar{\Phi}_{qij}(L)$ be the power series describing the response of variable X_{qt} to a shock X_{it} transiting through channel $X_{jt}j$. Then,*

1. *$\bar{\Phi}_{qij}(L)$ is given by*

$$\Phi_{qj}(L) \times \Phi_{ji}(L) = \Phi_{jj}(L) \times \bar{\Phi}_{qij}(L); \tag{5}$$

2. *$\bar{\Phi}_{qij}(L)$ is equivalent to the channel-specific response obtained using Sims and Zha (2006)'s channel shutoff method.*

Proof. The proof is developed in two steps. I first offer an implicit formulation for the IRFs, derive an expression for $\bar{\Phi}_{qij}(L)$ and show that it is equivalent to the expression obtained using Sims and Zha (2006)'s approach of shutting off the channel's response $\Phi_{ji}(L)$. I then proceed to prove that (5) holds.

Consider the AR(p) representation given in (64) with X_{it} a shock variable, X_{jt} a channel variable and X_{qt} a response variable. We are interested in the response $\bar{\Phi}_{qij}$ of X_{qt} to a shock on X_{it} that propagates via X_{jt}. To simplify the notation, let W_t be a $(n-2)\times 1$ vector grouping all the variables $\{X_{st} : s \neq j, s \neq q\}$. Note that the shock variable X_{it} is included in W_t. A general structural VAR representation is given by

$$AX_t = F(L)X_{t-1} + \eta_t \tag{64}$$

$$\begin{bmatrix} a_{ww} & a_{wj} & a_{wq} \\ a_{jw} & a_{jj} & a_{jq} \\ a_{qw} & a_{qj} & a_{qq} \end{bmatrix} \begin{bmatrix} W_t \\ X_{jt} \\ X_{qt} \end{bmatrix} = \begin{bmatrix} F_{ww}(L) & F_{wj}(L) & F_{wq}(L) \\ F_{jw}(L) & F_{jj}(L) & F_{jq}(L) \\ F_{qw}(L) & F_{qj}(L) & F_{qq}(L) \end{bmatrix} \begin{bmatrix} W_{t-1} \\ X_{j,t-1} \\ X_{q,t-1} \end{bmatrix} + \begin{bmatrix} \eta_{wt} \\ \eta_{jt} \\ \eta_{qt} \end{bmatrix}$$

where a_{ww}, $F_{ww}(L)$ are of dimension $(n-2)\times(n-2)$, $a_{wj}, a_{wq}, F_{wj}(L)$, $F_{wq}(L)$ are of dimension $(n-2)\times 1$ and $a_{jw}, a_{qw}, F_{jw}(L)$, $F_{qw}(L)$ are of dimension $1\times(n-2)$. Rearranging the first equation yields

$$\begin{aligned} \left(a_{ww} - LF_{ww}(L)\right) W_t &= \left(LF_{wq}(L) - a_{wq}\right) X_{qt} + \left(LF_{wj}(L) - a_{wj}\right) X_{jt} + \eta_{wt} \\ W_t &= \phi_{wq}(L)X_{qt} + \phi_{wj}(L)X_{jt} + \phi_{ww}(L)\eta_{wt}, \end{aligned}$$

with

$$\begin{aligned} \phi_{wq}(L) &\equiv \left(a_{ww} - LF_{ww}(L)\right)^{-1}\left(LF_{wq}(L) - a_{wq}\right); \\ \phi_{wj}(L) &\equiv \left(a_{ww} - LF_{ww}(L)\right)^{-1}\left(LF_{wj}(L) - a_{wj}\right); \end{aligned}$$

$$\phi_{ww}(L) \equiv \big(a_{ww} - LF_{ww}(L)\big)^{-1}$$

having respective dimensions $(n-2) \times 1$, $1 \times (n-2)$ and $(n-2) \times (n-2)$.

Taking the derivative with respect to η_{it} looking at horizon k and using Lemma 1 yields

$$\frac{dW_{t+k}}{d\eta_{it}} = \phi_{wq}(L)\frac{dX_{q,t+k}}{d\eta_{it}} + \phi_{wj}(L)\frac{dX_{j,t+k}}{d\eta_{it}} + \phi_{ww}(L)\frac{d\eta_{w,t+k}}{d\eta_{it}},$$
$$\Phi_{wi,k} = \phi_{wq}(L)\Phi_{qi,k} + \phi_{wj}(L)\Phi_{ji,k} + \phi_{ww}(L)\frac{d\eta_{w,t+k}}{d\eta_{it}}, \tag{65}$$

where $\Phi_{wi,k}$ is the vector of the responses $\{\Phi_{si,k} : s \neq j,q\}$. After recognizing that $\frac{d\eta_{q,t+k}}{d\eta_{it}} = \frac{d\eta_{q,t+k}}{d\eta_{jt}} = \frac{d\eta_{w,t+k}}{d\eta_{jt}} = 0$ for all horizons k, expressions for responses of X_{qt} to shocks on X_{it} and X_{jt}, and for the response of W_t to shocks on X_{jt} are obtained in the same manner with

$$\Phi_{qi,k} = \phi_{qj}(L)\Phi_{ji,k} + \phi_{qw}(L)\Phi_{wi,k} \tag{66}$$

$$\Phi_{qj,k} = \phi_{qj}(L)\Phi_{jj,k} + \phi_{qw}(L)\Phi_{wj,k} \tag{67}$$

$$\Phi_{wj,k} = \phi_{wq}(L)\Phi_{qj,k} + \phi_{wj}(L)\Phi_{jj,k}, \tag{68}$$

with $\phi_{qs}(L) \equiv \dfrac{LF_{qs}(L) - a_{qs}}{a_{qq} - LF_{qq}(L)}$ for all $s \in \{j, w\}$.

The total response of X_{qt} to a shock on X_{it} can be decomposed into components that do and do not involve the response of the channel variable X_{jt}. Substituting $\Phi_{wi,k}$ from (65) into (66) yields

$$\Phi_{qi,k} = \phi_{qj}(L)\Phi_{ji,k} + \phi_{qw}(L)\Big[\phi_{wq}(L)\Phi_{qi,k} + \phi_{wj}(L)\Phi_{ji,k} + \phi_{ww}(L)\frac{d\eta_{w,t+k}}{d\eta_{it}}\Big]$$
$$[1 - \phi_{qw}(L)\phi_{wq}(L)]\Phi_{qi,k} = [\phi_{qj}(L) + \phi_{qw}(L)\phi_{wj}(L)]\Phi_{ji,k} + \phi_{qw}(L)\phi_{ww}(L)\frac{d\eta_{w,t+k}}{d\eta_{it}}$$
$$\Phi_{qi,k} = \frac{\phi_{qj}(L) + \phi_{qw}(L)\phi_{wj}(L)}{1 - \phi_{qw}(L)\phi_{wq}(L)}\Phi_{ji,k} + \frac{\phi_{qw}(L)\phi_{ww}(L)}{1 - \phi_{qw}(L)\phi_{wq}(L)}\frac{d\eta_{w,t+k}}{d\eta_{it}}. \tag{69}$$

The component of X_{qt}'s response to X_{it} propagating via X_{jt} is thus given by

$$\bar{\Phi}_{qij,k} = \psi(L)\Phi_{ji,k} \tag{70}$$

where $\psi(L) = \frac{\phi_{qj}(L)+\phi_{qw}(L)\phi_{wj}(L)}{1-\phi_{qw}(L)\phi_{wq}}$. This result will be useful in the second part of the proof to prove that (5) holds.

Sims and Zha (2006) procedure of shutting off the response of the channel variable X_{jt} to the shock on X_{it} is equivalent to setting $\Phi_{ji,k} = 0$ for all k. The channel specific response $\bar{\Phi}_{qij,k}$ is thus equivalent to that obtained by Sims and Zha (2006) and the second part of the proposition is established.

The first part of the proposition remains to be demonstrated and is proven hereinafter. Let $C(L)$ be the power series obtained by convoluting the power series $\Phi_{qj}(L)$ and $\Phi_{ji}(L)$ describing respectively the response of X_{qt} to a shock on X_{jt} and the response of X_{jt} to a shock on X_{it}.

$$\begin{aligned}
C(L) &= \Phi_{qj}(L) \times \Phi_{ji}(L) \\
C(L) &= \Big[\sum_{k=0}^{\infty} \frac{dX_{q,t+k}}{d\eta_{jt}} L^k\Big] \times \Big[\sum_{k=0}^{\infty} \frac{dX_{j,t+k}}{d\eta_{it}} L^k\Big] \\
C(L) &= \Big[\sum_{k=0}^{\infty} \Phi_{qj,k} L^k\Big] \times \Big[\sum_{k=0}^{\infty} \Phi_{ji,k} L^k\Big].
\end{aligned} \tag{71}$$

An expression for $\Phi_{qj,k}$ is given in equation (67). Substituting $\Phi_{wj,k}$ from (68) into (67) gives

$$\begin{aligned}
\Phi_{qj,k} &= \phi_{qj}(L)\Phi_{jj,k} + \phi_{qw}(L)\Phi_{wj,k} \\
\Phi_{qj,k} &= \phi_{qj}(L)\Phi_{jj,k} + \phi_{qw}(L)\Big[\phi_{wq}(L)\Phi_{qj,k} + \phi_{wj}(L)\Phi_{jj,k}\Big].
\end{aligned} \tag{72}$$

And after rearranging, we have

$$\begin{aligned}
\Phi_{qj,k} &= \frac{\phi_{qj}(L) + \phi_{qw}(L)\phi_{wj}(L)}{1 - \phi_{qw}(L)\phi_{wq}(L)}\Phi_{jj,k} \\
\Phi_{qj,k} &= \psi(L)\Phi_{jj,k},
\end{aligned} \tag{73}$$

where $\psi(L)$ is defined as above. Substituting in (71) yields

$$C(L) = \Big[\sum_{k=0}^{\infty} \psi(L)\Phi_{jj,k}L^k\Big] \times \Big[\sum_{k=0}^{\infty} \Phi_{ji,k}L^k\Big]$$
$$C(L) = \Big[\sum_{k=0}^{\infty} \Phi_{jj,k}L^k\Big] \times \Big[\sum_{k=0}^{\infty} \psi(L)\Phi_{ji,k}L^k\Big]$$
$$C(L) = \Phi_{jj}(L) \times \Big[\psi(L)\Phi_{ji}(L)\Big].$$

Substituting from (70), we have

$$C(L) = \Phi_{jj}(L) \times \bar{\Phi}_{qij}(L)$$

and (5) holds ■