# Inequality of Opportunity, Inequality of Effort, and Innovation*

Alessandro Spiganti†

## Abstract

How are inequality of opportunity, innovation, and inequality of outcomes interlinked? In this paper, I present an occupational choice model where agents differing in public wealth and private ability choose between working for a wage or becoming inventors. In a Schumpeterian fashion, inventors are responsible for a random sequence of quality-improving innovations. I show that occupational decisions depend on both the initial wealth distribution and the current degree of technological progress. Since these are both endogenous, I extend the model dynamically by assuming that agents are periodically replaced by a new generation, to whom they leave bequests. Empirically, I show that there seems to be a negative long-term relationship between the inequality of opportunity of a US state and the number of patents per capita filed in that state, but a positive relationship between patenting activity and inequality of effort.

*Keywords:* Innovation; Inequality; Occupational Choices; Bequests; Theil's index
*JEL Classification:* D15, D53, D58, D82, H23, O31
*Word count:* approximately 6,000

*Thanks go to Catherine Bobtcheff, Andrew Clausen, Philipp Kircher, Salvatore Lo Bello, Domenico Menicucci, John H. Moore, Francisco Queirós, and seminar participants at SMYE 2019, and EUI.
†European University Institute, San Domenico di Fiesole, I-50014, Italy. Email: Alessandro.Spiganti@EUI.eu

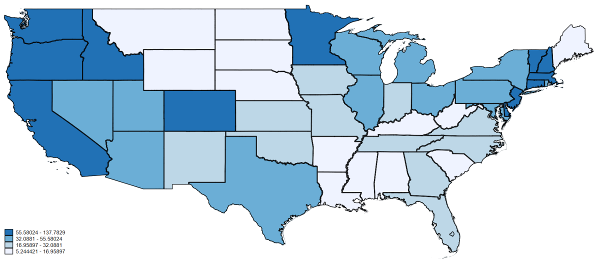
## I. Introduction

It is widely recognised that OECD countries have experienced a sharp increase in inequality in the past 50 years. For example, the average income of the richest 10% of the population is about nine times that of the poorest 10% across the OECD countries, up seven-fold from 25 years ago. In the years since the financial crisis, concerns about this increase have entered the political and economic mainstream, sparkling a new wave of economic literature on the macroeconomic effects of inequality. In most cases, the question boils down to this: is rising inequality good or bad for growth? In this article, I study the relationship between inequality and innovation, one of the major drivers of economic growth, and argue that to study this relationship it is important to separate inequality into two components: inequality of opportunity, stemming from factors beyond an individual's control, and inequality of effort, caused by individual responsible choices. Figure 1, for example, maps the average quality-adjusted number of patents filed in a given US state over the period 1976 - 2006, a proxy of inequality of opportunity, and one for inequality of effort. By comparing the different panels, a set of anecdotal evidence begins to emerge. First, I see that states may have relatively different measures of inequality. Second, US states that are traditionally considered as the most innovative,[1] and indeed have a relatively high average number of patents per capita, like California, Washington, Massachusetts, and Colorado, tend to have low levels of inequality of opportunity and high levels of inequality of effort. Finally, less innovative states, like Mississippi, Louisiana, West Virginia, Tennessee, and Arkansas, rank relatively high in inequality, especially of opportunity.

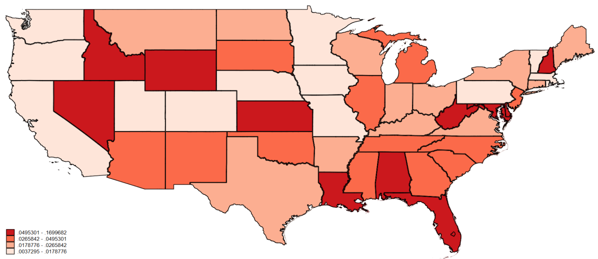
Intuitively, a more unequal reward structure could provide the financial incentives necessary for risky activities, like innovation; conversely, a more equal playing field could allow more agents to realise their innovative potential. In this paper, I highlight the following caveat: a more unequal reward structure could hinder the realisation of a more equal future playing field, if, for example, the fortune at birth is linked with the fortune of the parents. To what extent then is there an optimal degree of inequality for innovation? Can countries be both innovative and have equal societies?

To answer these questions, I first construct a theoretical model with three components: (i) heterogeneous agents, (ii) occupational choices, and (iii) Schumpeterian innovations. Agents are heterogeneous in unobservable talent and observable wealth, and must choose between working for a wage or becoming inventors. Incentivised by the prospects of monopoly rents, inventors create better machines for production. Long-run growth results from creative destruction: these productivity-improving innovations replace old machines

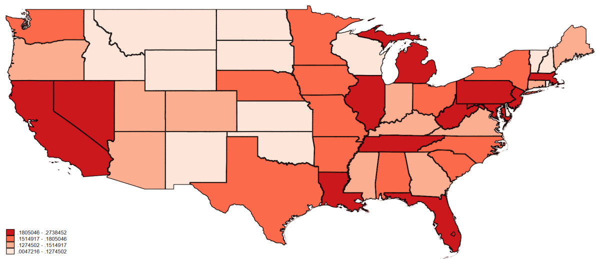
[1] See, for example, the article by Adam McCann published on WalletHub on 18 March 2019, "Most & Least Innovative States" (available online at `https://wallethub.com/edu/most-innovative-states/31890/`).



(a) Quality-Adjusted Number of Patents



(b) Inequality of Opportunity



(c) Inequality of Effort

Figure 1: Patents and Inequality by US State

*Notes.* The fist panel refers to the average number of patents granted by the United States Patent and Trademark Office to a patent inventor resident in the state per hundred thousand residents, and weighted by the number of citations received within 5 years of the application date. Data are elaborated by the author from Aghion *et al.* (2019) and refer to the period 1976 - 2006. The last two panels are 1968 - 2017 averages of the decompositions of the Theil's $T$ index calculated by the author using data from the Panel Study of Income Dynamics.

and feature positive intertemporal spillovers.

In first best, the most talented agents would be employed in the innovation sector, to maximise the growth rate of the economy. Unfortunately, becoming an innovator requires an initial investment and, since talent is unobservable, credit market frictions may prevent an efficient allocation. In equilibrium, adverse selection prevents the poor but talented agents becoming innovators, who are displaced by relatively wealthier but untalented agents. I show that the seriousness of the adverse selection problem depends on both the current level of innovation, which influences the profits from innovating, and the initial wealth distribution. Since these are endogenous and influence one another over time, I extend the model dynamically, by assuming that old agents are periodically replaced by a new generation, to whom they transfer wealth. I show that the initial wealth distributions have long-run effects and that policies aimed at increasing the number of innovations in a given period may have unintended negative consequences for long-run growth.

Then, I present an empirical analysis where I investigate the relationship between inequality and innovation at the US state level. I measure innovation using patent and citations data from the United States Patent and Trademark Office, going back to 1976. I use labour income data covering the period 1968 - 2017 from the Panel Study of Income Dynamics to calculate total inequality rate at the state level. I employ widely-used technique in the inequality of opportunity literature to separate total inequality into inequality between socio-economic groups (classified by race and parental education) and inequality within groups: the first component represents a proxy for inequality due to circumstances beyond the individual's control (i.e. inequality of opportunity), whereas the latter is a proxy for inequality due to individual choices (i.e. inequality of effort). I find that, whereas innovation is positively correlated with total inequality, this seems to be the result of a negative correlation with inequality of opportunity and a positive correlation with inequality of effort.

The remainder of this paper is organised as follows. Section II quickly reviews previous literature. Section III presents the theoretical model, and Section IV finds its partial equilibrium. Section V outlines the empirical analysis. Finally, Section VI concludes.

## II. Previous Literature

This paper bridges the Schumpeterian growth theory literature pioneered by Aghion and Howitt (1992) with the literature on the effect of misallocation on growth (see Murphy *et al.*, 1989, Banerjee and Newman, 1993, Galor and Zeira, 1993, for some seminal contributions). More broadly, this paper is also related to the literature on the consequences of occupational choices for inequality (Kambourov and Manovskii, 2009), on innovation incentives (e.g. Holmström, 1989, Aghion and Tirole, 1994, Manso, 2011, Spiganti, forthcoming), and on occupational persistence across generations (e.g. Caselli and Gennaioli, 2013, Lo Bello and Morchio, 2016).

More specifically, this paper belongs to a growing literature on the relationship between inequality and innovation.[2] Recently, Aghion *et al.* (2019) and Jones and Kim (forthcoming) have built Schumpeterian models that link the dynamics of top income inequality to innovation, and showed that creative destruction makes growth more inclusive. Acemoglu *et al.* (2017) positively link innovative activities of an economy to a more unequal reward structure, whereas Spiganti (2018) studies occupational choices into innovative activities under wealth inequality and heterogeneous innovative talent. Differently from these papers, I focus on both wealth and income inequality, and the feedback effect between them thanks to the presence of intergenerational linkages.

The long term effect of the misallocation of talent in innovative activities when there are intergenerational linkages is also the focus of Jaimovich (2011) and Celik (2018). However, there are several differences between our papers. First, in Celik's (2018) quantitative model, inventors are skilled workers employed by firms for a fixed wage, whereas in Jaimovich (2011), horizontal innovation occurs when agents open up new sectors match-

[2] I stress that in this paper I study neither the effect on income inequality of the introduction of new technologies (see Violante, 2008, for a brief survey on skill-biased technological change), nor the effect of inequality on the incentives to innovate through demand composition (Murphy *et al.*, 1989, Zweimüller, 2000, Foellmi and Zweimüller, 2006, 2017, Hatipoğlu, 2012).

ing their intrinsic skills. In this paper, innovation is vertical and Schumpeterian: inventors are entrepreneurs who are willing to face the risk of failure to discover better vintages of existing machines and pursue monopoly rents. This allows me to study how the occupational choice into innovation is shaped by different reward structures. Second, in Celik (2018) everyone would like an innovative job, as it pays exogenously better than routine jobs, but the number of training opportunities necessary to become skilled is scarce and subject to a tournament mechanism; in Jaimovich (2011), there is no occupational choice as everyone is an entrepreneur, but adverse selection may prevent credit flowing to the most productive sectors (i.e. with a better match between an entrepreneur's skill and a sector's characteristics). Here, observable wealth is used by banks to screen different borrowers: as a consequence, endogenous wealth classes arise in equilibrium, each associated with different occupational choices.[3] Since the wealth distribution affects the composition of the wealth classes, and the occupational choices of the agents, it affects the resulting growth rate of the economy. Moreover, this indirectly affects the new wealth distribution. This allows me to study how the number of innovators and their average quality change vis-à-vis the state of the economy.

Empirically, there is a very recent and flowering literature on the relationship between inequality and innovation. For example, Akcigit *et al.* (2017), Aghion *et al.* (2018), Celik (2018), and Bell *et al.* (forthcoming) merge individual income data with individual patenting data and find a positive relationship between parental resources and the probability of becoming an inventor. Conversely, Aghion *et al.* (2019) find a positive effect of patenting on top income inequality, using a US state level panel. In this paper, I follow techniques that are widely used in the inequality of opportunity literature to decompose total inequality at the US state level into inequality of opportunity and inequality of effort (see, for example, Rodríguez, 2008, Ferreira and Gignoux, 2011, Marrero and Rodríguez, 2013, Liao, 2016a). This allows me to consider the relationship between innovation and both components, whereas the above papers consider only one.

## III. A Theoretical Model

Time is discrete and infinite, $t = 1, 2, \dots, \infty$. In any given period $t$, there is a continuum of one-period lived agents of mass one, indexed by $h$, with the same instantaneous utility function,

$$u(c_{t,h}, b_{t,h}) = c_{t,h}^{1-\delta} b_{t,h}^{\delta}, \tag{1}$$

[3] On the modelling side, the interplay between adverse selection and occupational choices in this paper is reminiscent of e.g. Grüner (2003), Ghatak *et al.* (2007), Inci (2013), and Spiganti (2018). Differently from these papers, I extend this framework to a dynamic setting and embed it into a Schumpeterian model of innovation.

where $\delta \in (0,1)$, $c_{t,h}$ is consumption of the final good, and $b_{t,h}$ is bequest in period $t$.[4]

Agents are heterogeneous in two dimensions. First, they differ in their innovative ability $a$: in any period $t$, a proportion $\lambda \in (0,1)$ of agents is talented, the remaining proportion $1-\lambda$ is untalented. Second, agents differ in their wealth endowment, $A_t$, which is distributed according to the continuously differentiable cumulative distribution function $\Phi_t(A)$, whose probability density function is $\phi_t(A)$. Let $\bar{A}_t = \int_0^\infty A_t d\Phi_t(A)$ be average (and total) wealth in $t$. For simplicity, I assume that ability and wealth are uncorrelated: this implies that abilities are also intergenerationally uncorrelated and that there is an equal proportion of talented and untalented agents for every wealth level. At the beginning of their life, agents receive their wealth in the form of a bequest from their parent.

## III.1 Final Good Production

Agents consume an homogeneous final good, $y_t$. This is produced competitively by a representative firm combining unskilled labour and a continuum of machines indexed on the interval $[0,1]$ according to

$$y_t = f(l_t, x_{t,m}) = l_t^{1-\alpha} \int_0^1 Q_{t,m}^{1-\alpha} x_{t,m}^{\alpha} dm, \tag{2}$$

where $\alpha \in (0,1)$, $l_t$ is labour, $Q_{t,m}$ is the quality of machine of type $m$ used, and $x_{t,m}$ is the quantity of this machine.[5] Let $Q_t \equiv \int_0^1 Q_{t,m} dm$ be the average quality of the machines, an aggregate quality index of the economy.

The profit-maximization problem of the final good producer is

$$\pi_t\left(p_t; w_t, r_{t,m}\right) = \max_{l_t, \{x_{t,m}\}_{m=0}^1 \geq 0} p_t f(l_t, x_{t,m}) - w_t l_t - \int_0^1 r_{t,m} x_{t,m} dm, \tag{3}$$

where $p_t$ is the price of the final good, $w_t$ is the wage rate, and $r_{t,m}$ is the price of machine of type $m$ used. The first order conditions (henceforth, FOCs) are

$$(1-\alpha) p_t l_t^{-\alpha} \int_0^1 Q_{t,m}^{1-\alpha} x_{t,m}^{\alpha} dm = w_t \tag{4a}$$

$$\alpha p_t l_t^{1-\alpha} Q_{t,m}^{1-\alpha} x_{t,m}^{\alpha-1} = r_{t,m} \tag{4b}$$

[4] In line with the "warm glow" or "joy of giving" literature that follows from Andreoni (1989, 1990), I assume that bequests, rather than offspring's utility, enter the utility function directly. Under this assumption, utility is linear in end-of-period wealth, and this makes the model more tractable (see e.g. Banerjee and Newman, 1993, Jaimovich, 2011, for an identical assumption).

[5] Similar formulations of this multisector Schumpeterian model of endogenous growth (i.e. where growth is generated by a random sequence of vertical improvements) appear in Aghion and Howitt (2009, Ch. 4) and Acemoglu *et al.* (2012). Note that there is nothing of importance lost by having $l_t$ and $Q_{t,m}$ raised to the same power (see Aghion and Howitt, 2009, Ch. 4, problem 2), and that I are implicitly assuming that production uses only the highest quality machine for each type.

and thus the following demand curves are obtained (for ease of reading, I are ignoring that some of the right-hand side variables are policy functions):

$$l_t(p_t; w_t, r_{t,m}) = \left( \frac{p_t(1-\alpha)}{w_t} \int_0^1 Q_{t,m}^{1-\alpha} x_{t,m}^{\alpha} dm \right)^{\frac{1}{\alpha}} \tag{5a}$$

$$x_{t,m}(p_t; w_t, r_{t,m}) = \left( \frac{\alpha p_t}{r_{t,m}} \right)^{\frac{1}{1-\alpha}} Q_{t,m} l_t. \tag{5b}$$

It will prove useful to notice that these demands are iso-elastic with elasticities given by $-1/\alpha$ and $-1/(1-\alpha)$ for labour and machine, respectively.

## III.2 Innovation

Innovation in each machine takes place as follows. Becoming an inventor requires an exogenous sunk cost of $I_t$. An innovator is then matched randomly with one machine (one to one, no congestion). Producing one unit of any machine costs $\psi$ units of final good.

Innovation is stochastic, with probabilities of success depending on the innovative talent of the agent. Low-type agents always result successful with probability $\rho_L$. High-ability individuals can raise the probability of success to $\rho_H$ by working hard, but this comes at a positive cost $e$, which is measured in monetary units.[6] Hereafter, effort-exerting talented agents are denoted by $H$ (mnemonic for high-ability), whereas untalented and shirking talented agents are denoted by $L$ (for low-ability). The ability of the agents and their effort level are known only by them, but the distribution of talent in every wealth level is public information.

In case of success, an innovator increases the quality of the machine from $Q_{t,m}$ to $(1+\gamma)Q_{t,m} > Q_{t,m}$ and, in line with the endogenous technical change literature, becomes the sole producer of the machine $m$. The profit-maximization problem of the inventor of a new machine $m$ is

$$\max_{r_{t,m}^M, X_{t,m}^M \geq 0} (r_{t,m}^M - \psi) X_{t,m}^M \qquad \text{s.t. } X_{t,m}^M \geq x_{t,m} \text{ in (5b)}, \tag{6}$$

where $r_{t,m}^M$ and $X_{t,m}^M$ are the price and quantity supplied of the monopolistically-produced machine $m$ in $t$. Since demand is iso-elastic, the monopoly price is a constant mark-up over marginal cost, $r_{t,m}^M = \psi/\alpha$,[7] and the equilibrium demand function for monopolistically-

[6]The talent distribution in the population can thus be thought of as a distribution of the cost of effort, which is prohibitively high for untalented individuals. Intuitively, everyone in this economy is born untalented: some individuals have the potential to undertake some costly activity to increase their talent, whereas those that remain lack the natural ability. Similarly to Grüner (2003), Inci (2013), and Spiganti (2018), moral hazard is necessary to have some poor talented workers in equilibrium.

[7]I are implicitly assuming, for simplicity, that innovation is drastic, in the sense of Tirole (1988): the monopolist can charge any price she wants without fearing entry from potential competitors.

produced machines, $x^M_{t,m}$, is

$$x^M_{t,m} = \left(\frac{\alpha^2 p_t}{\psi}\right)^{\frac{1}{1-\alpha}} Q_{t,m} l_t. \tag{7}$$

Here, I make a further simplifying assumption. Similarly to Aghion and Howitt (2009, Ch. 6), I assume that the starting quality for any given machine $m$ at date $t$ has the average quality parameter $Q_{t-1}$ across all machines last period, rather than the quality parameter $Q_{t-1,m}$ of that machine last period.[8] Therefore, an innovator that is successful in inventing a new machine, would profit

$$\varpi_t\left(p_t; w_t; Q_{t-1}\right) \equiv \left(\frac{\psi}{\alpha} - \psi\right)\left(\frac{\alpha^2 p_t}{\psi}\right)^{\frac{1}{1-\alpha}} (1+\gamma) Q_{t-1} l_t \tag{8}$$

from selling the machine.

With probability $1-\rho_i$, $\forall i = \{H, L\}$, the innovation does not materialise. In such case, the old machine is produced competitively. Let $X^C_{t,m}$ be the quantity of the competitively-produced machine $m$ in $t$. Since the unsuccessful innovator prices the machine at the marginal cost, $r^C_{t,m} = \psi$, the equilibrium demand function for competitively-produced machines is

$$x^C_{t,m} = \left(\frac{\alpha p_t}{\psi}\right)^{\frac{1}{1-\alpha}} Q_{t,m} l_t. \tag{9}$$

The unsuccessful innovator breaks even.

## III.3 Credit Contracts

In each period, there are several banks competing à la Bertrand, each owned equally by all agents. Workers deposit their wealth in the banks for a risk-free rate of return, $R_t$: an investment of one unit in $t$ yields a return of $R_t$ units at the end of the period. All agents take this rate of return as given when making their occupational choices. Banks use these deposits to lend money to innovators that ask for it: without loss of generality, I assume that an agent with wealth $A_t$ only borrows up to $I_t - A_t > 0$ to finance the set-up cost; conversely, rich innovators deposit $A_t - I_t > 0$.[9] In the next sections, unless otherwise specified, I focus on the credit constrained agents when deriving the optimal contracts, as unconstrained agents are free to take their first-best choice.

Banks observe the wealth of the borrowers, and whether they succeeded or not in

[8] I make this assumption to avoid further complications arising from having to include the quality of the machine in the optimal contract derived below.

[9] It is well-known, see e.g. DeMeza and Webb (1987), that there must be maximum self-finance in equilibrium, because this comes with better terms than borrowing for high ability agents (thus, if there are agents who are not using their entire wealth, they must be untalented).

discovering a new machine, but ability is unobservable.[10] They take prices, including the riskless rate of return, as given, and can offer a distinct menu of contract for every wealth level. Banks hold the same beliefs, which they form simultaneously, about how agents decide when offered a given menu of contracts. This menu of contracts consists of a repayment schedule, given the factor prices and qualities of the machines, contingent on the outcome of the innovation process and the announced type. I assume limited liability protects the agents, in the sense that an innovator cannot be left with negative end-of-period payoff. As a consequence, and since in the failure state innovators break even, they will be able to pay back a positive amount only in the case of success.[11] A loan contract offered by a given bank then takes the following form (since it does not generate confusion, I shall drop the subscript indicating a given bank and $t$):

$$\boldsymbol{\sigma}(A,\Omega)=\begin{bmatrix}\sigma_H(A,\Omega)\\ \sigma_L(A,\Omega)\end{bmatrix}=\begin{bmatrix}D_H^S(A,\Omega)\\ D_L^S(A,\Omega)\end{bmatrix}, \tag{10}$$

where $\sigma_i$ is the contract designed for the $i$-type agents with wealth $A$ and $\Omega$ is a vector of prices and average quality of the machine at $t$ (i.e. the *state* of the economy). These contracts set the repayments to the bank by the $i$-type agent in the success state, $D_i^S$.

Therefore, with probability $\rho_i$, the innovator is successful in inventing a new machine, and thus profit $\varpi(\Omega)$ from selling the machine. She will then pay $D_i^S(A,\Omega)$ to the bank. Let the subsequent realised net payoff of an $i$-type innovator in the success state be given by $V_i^S$; limited liability implies $V_i^S \geq 0$. Conversely, an unsuccessful innovator would break even from the production of machines, and the financiers would recover zero income. Thus, the expected payoff of an innovator is $\mathcal{V}_i(A,\Omega) \equiv \rho_i V_i^S - e_i$, $\forall i = \{H, L\}$, where $e_H = e$ and $e_L = 0$. Conversely, the payoff of an agent who becomes a worker is $\mathcal{W}_i(A,\Omega) \equiv w + RA, \forall i = \{H, L\}$: I call this payoff "the outside option" to innovation.

## IV. Partial Equilibrium Analysis

In this section, I first analyse the static counterpart of our model: I focus on a given period, and thus the ability and wealth distributions, as well as the average quality of the machines, are given. I thus derive the set of credit contracts offered by banks, and the optimal occupational choices of the agents. Later, I study the dynamic evolution of the economy. Throughout this section, I abstract from general equilibrium effects in the credit and labour markets, i.e. I take the risk-free interest rate and the wage rate as

[10]Since abilities are intergenerationally uncorrelated, parents' historical outcomes provide no useful information to the banks.

[11]This means that the repayment in case of failure cannot be positive, but, in principle, it may be the case that banks offer money to unsuccessful innovators. Given risk-neutrality, however, imposing the repayments to be zero in the failure state is without loss of generality. Appendix A.2 presents the proofs without imposing this.

exogenously given.[12]

I take the standard assumption that talented innovation is efficient, whereas untalented innovation is not. This means that, if agents could self-finance completely, only talented agents would find it profitable to enter the innovation sector. However, untalented agents may still find it profitable to become innovators if cross-subsidised by talented agents. This is formalised as follows,

**Assumption 1** (Static Efficiency). $\rho_H \varpi(\Omega) - e > w + RI > \rho_L \varpi(\Omega) > w + RI(\rho_L/\bar{\rho})$,

where $\bar{\rho} = \lambda\rho_H + (1-\lambda)\rho_L$ is the Bayesian probability of success of a random applicant.

I impose a Bertrand-Nash equilibrium concept in the static framework. As it is well-known from Rothschild and Stiglitz (1976), this may lead to non-existence of a competitive screening equilibrium. I circumvent this by restricting the set of feasible contracts to loan contracts only, i.e. non-negative repayments made by entrepreneurs to the banks.[13]

**Equilibrium Concept.** *Assume banks are Bertrand-Nash players following pure strategies, offering loan contracts, and paying an interest $R$ on deposits, which they take as given. A static equilibrium consists of choices for individuals $(c_h^\star, b_h^\star)$, the final good producer $(l^\star, \{x_m^\star\}_{m=0}^1)$, and the innovators $(\{r_m^\star, X_m^\star\}_{m=0}^1)$; prices $(w^\star, p^\star)$; profits for banks, final good producer, and innovators; sector allocations and effort decisions; and an individually rational and incentive compatible menu of contracts for each wealth class, such that: (i) banks earn non-negative profits at every wealth level, (ii) the machine and final good producers maximise their profits, (iii) the menu of contract is a Bertrand-Nash equilibrium, (iv) the machines and final good markets clear, (v) individuals choose the occupation that maximises their expected end-of-period wealth, and (vi) talented individuals choose innovation if indifferent between the two occupations, whereas untalented individuals choose wage-earning when indifferent.*

[12] For example, this would be the case if our economy is small, and with access to perfect international capital and labour markets. Financial intermediaries (depositors) would thus be able to draw (deposit) liquid funds from (in) an international credit market, with a perfectly elastic supply and demand at the international rate of return $R$. Firms would be able to hire workers with a perfectly elastic supply. A sketch of the general equilibrium counterpart is available on request.

[13] Practically, this means that banks cannot lure in additional depositors by increasing the interest rates they offer to lenders. As explained below, this restriction means that banks will make positive profits on the contracts offered to particular wealth classes, like in Jaimovich (2011). This is, however, a Bertrand-Nash equilibrium as there are no profitable deviations in the set of feasible contracts. One could, alternatively, enlarge the set of feasible contracts and then either impose a Bertrand-Wilson's (1977) anticipatory equilibrium concept, as in Inci (2013), or allow banks two rounds of play, like in Hellwig (1987). In any case, the qualitative results of our model do not change. See Appendix A.2 for more detail.

## IV.1 The Equilibrium Under Full Information

Suppose information about agents' skills were complete, so that in equilibrium, banks would charge an interest rate that accurately reflects an agent's intrinsic risk of failure. Since there cannot be any cross-subsidisation in an equilibrium without adverse selection, talented individuals become innovators, facing a rate of return equal to $R_t/\rho_H$, and an expected end-of-period wealth of $\mathcal{V}_t^{FB}(A_t, \Omega_t) = \rho_H \varpi(\Omega_t) - R_t(I_t - A_t) - e$. Conversely, untalented agents become workers, with an end-of-period wealth of $\mathcal{W}_t^{FB}(A_t, \Omega_t) = w_t + R_t A_t$. Note that, by Assumption 1, $\mathcal{V}_t^{FB}(A_t, \Omega_t) - \mathcal{W}_t^{FB}(A_t, \Omega_t) > 0$: as we will see below, this talent premium reaches its maximum value when information asymmetries are absent.

Under full information, the average probability of success of the innovators is equal to $\rho_H$, and there are $\lambda$ innovators each period. Thus, the expected number of successful innovations in any given period is given by $\rho_H \lambda$. Conversely, the expected number of unsuccessful innovations is $1-\rho_H\lambda$. Since innovations increase the quality of the machines to $(1+\gamma)Q_{t-1}$, whereas failure leaves the quality equal to $Q_{t-1}$, the average quality of the machines increases over time with a growth rate of $g^{FB} = \gamma \rho_H \lambda$.[14]

## IV.2 Static Equilibrium

Below, I intuitively derive the static partial equilibrium of the model, i.e. the contracts offered and the subsequent occupational choices of the agents, taking the state of the economy and the wealth distribution as given. The proofs are formally given in Appendix A.2. For readability, I drop the time subscript.

Following the literature on adverse selection, one should expect two types of equilibrium contracts: pooling, in which types remain undistinguishable, or separating, in which types reveal their unobservable ability by selecting different terms. Here, one can easily exclude that there exists a separating loan contract, for a given wealth class, such that both talented and untalented agents become innovators. To see this, consider the following zero-profit conditions from separating contracts,

$$\rho_H \left(\varpi\left(\Omega\right) - V_H^S\right) = R(I - A) \tag{11a}$$

$$\rho_L \left(\varpi\left(\Omega\right) - V_L^S\right) = R(I - A), \tag{11b}$$

where the first line refers to talented agents, and the second one to untalented agents with wealth $A$. The implied levels of $\mathcal{V}_H$ and $\mathcal{V}_L$ suggest that this menu of contract cannot be

[14]Note that, in first best, all machines would be produced competitively, whereas here there is still an inefficiency due to the presence of monopoly rights granted to the successful innovators. The resulting underutilisation of machines can easily be corrected with a subsidy in the use of (new vintages of) machines, such that their net price is identical to the marginal cost.

incentive compatible, as the untalented agents would always prefer the contract designed for the talented innovators.

Hence, an equilibrium contract must be either a pooling contract or a separating contract that only the talented type accepts.[15] In a zero-profit pooling contract, the repayment of a random borrower with wealth $A$ in the success state is given by $D^S = R(I-A)/\bar{\rho}$. For a given state of the world $\Omega$, an $i$-type agent would accept this contract if her participation constraint is satisfied,

$$\rho_i \left( \varpi \left( \Omega \right) - \frac{R \left( I - A \right)}{\bar{\rho}} \right) - e_i \geq w + RA. \tag{12}$$

At the same time, talented agents would be willing to exert effort only if

$$\rho_H \left( \varpi \left( \Omega \right) - \frac{R \left( I - A \right)}{\bar{\rho}} \right) - e_H \geq \rho_L \left( \varpi \left( \Omega \right) - \frac{R \left( I - A \right)}{\bar{\rho}} \right). \tag{13}$$

Solving these for $A$ reveals that talented agents exert effort with a pooling contract if their wealth is greater than a threshold $A_e$, and they enter the innovation sector if their wealth is greater than a threshold $A_H$. Conversely, untalented agents become innovators only if their wealth is lower than a threshold $A_L$. These thresholds are given by, respectively,

$$A_e \equiv I + \frac{\bar{\rho} \left( e - \left( \rho_H - \rho_L \right) \varpi \left( \Omega \right) \right)}{R \left( \rho_H - \rho_L \right)} \tag{14a}$$

$$A_H \equiv \frac{\bar{\rho}(w+e) + \rho_H \left( RI - \bar{\rho} \varpi \left( \Omega \right) \right)}{R \left( \rho_H - \bar{\rho} \right)} \tag{14b}$$

$$A_L \equiv \frac{\rho_L \left( \bar{\rho} \varpi \left( \Omega \right) - RI \right) - \bar{\rho} w}{R \left( \bar{\rho} - \rho_L \right)}. \tag{14c}$$

As common in this literature, talented agents are only willing to accept a pooling contract if the amount they need to borrow is small: indeed, since banks underestimate their probability of success, they are obliged to subsidise the untalented agents. Conversely, untalented agents only accept pooling contracts if they can enjoy large cross-subsidies.

Note that by Assumption 1, $A_L > 0$. Throughout this paper, I focus on the most interesting case by further assuming that $A_L > A_e > 0$: this ensures that there is both adverse selection and some poor talented workers in equilibrium. This in turn implies that $A_e > A_H$. Wherever $A_H$ lies, for agents with $A < A_e$, this pooling contract cannot be offered: indeed, the average probability of success of the innovators in this class would be $\rho_L$, resulting in negative profits for the banks. As a consequence, the only contract that can be offered for these wealth levels is the one on the zero-profit condition from

[15] It is easy to prove, given the probabilities of success, that it can never be the case that a loan contract, for a given wealth level, attracts the untalented agent but not the talented agent.

untalented innovation, with $D^S = R(I-A)/\rho_L$. Since everyone is treated as untalented, by Assumption 1 all agents in this wealth bracket prefer to become workers.

The last possibility is that, for a given wealth class, only the effort-exerting talented agents enter the innovation sector. A putative separating contract on the zero-profit condition entails $D_H^S = R(I-A)/\rho_H$: I refer to this contract as "the zero-profit separating contract". Obviously, this contract can be offered only if an untalented agent with the same wealth does not have any incentive to imitate the talented agent, i.e. if

$$w + RA \geq \rho_L \left( \varpi(\Omega) - \frac{R(I-A)}{\rho_H} \right). \tag{15}$$

This condition requires that her initial wealth is higher than a threshold $A_{HH}$ given by

$$A_{HH} \equiv \frac{\rho_L(\rho_H \varpi(\Omega) - RI) - \rho_H w}{R(\rho_H - \rho_L)}, \tag{16}$$

where $A_{HH} < I$ by Assumption 1.

Since for $A \in (A_L, A_{HH})$ the zero-profit separating contract above does not satisfy the incentive compatibility constraint of an untalented agent with identical wealth, these talented agents will have to receive a different contract. The solution involves raising the interest rate demanded of the talented agents in such a way that makes the untalented agents indifferent between entering the innovation sector and becoming workers. This is achieved by imposing $\rho_L\left(\varpi(\Omega) - D_H^S\right) = w + RA$, or, equivalently, $D_H^S = \varpi(\Omega) - (w + RA)/\rho_L$. I refer to this contract as "the profitable separating contract".

Given the set of contracts offered to each wealth class, it is easy to derive the resulting occupational choices of the agents. Proposition 1 outlines the static partial equilibrium that ensues.

**Proposition 1** (Static partial equilibrium)**.** *Banks offer contracts on the zero-profit condition from untalented innovation to agents with wealth in $[0, A_e]$, pooling contracts to agents in $[A_e, A_L]$, profitable separating contracts to agents in $[A_L, A_{HH}]$, and zero-profit separating contracts to agents in $[A_{HH}, I]$. This is associated with the following occupational choices: all agents in $[0, A_e]$ become workers, all agents in $[A_e, A_L]$ become innovators, talented agents with $A \geq A_L$ become innovators, whereas their untalented counterparts become workers.*

**IV.2.1 Wealth Classes.** Proposition 1 underlines that we can split the population into different pools of borrowers depending on their wealth level. Indeed, the contractual structure of the lending market endogenously introduces four wealth classes, that I label working, lemons, rich, and unconstrained.

The working-class agents have wealth between $[0, A_e]$. Given the size of the loan that they would need, talented agents do not apply for loans, and thus the only offer banks can make is an interest rate of $R/\rho_L$. As a consequence, every agent in this class becomes a worker, with an end-of-period wealth of $w + RA$. The lemons-class agents have wealth in $[A_e, A_L]$. Banks offer only pooling contracts, with an interest rate of $R/\bar{\rho}$, and both types of agents become innovators. The expected end-of-period wealth of an $i$-type agent in this class is $\rho_i \left(\varpi - R(I - A)/\bar{\rho}\right) - e_i$. The rich-class agents have wealth in $[A_L, A_{HH}]$. Banks can offer separating contracts to agents in this class, but with an interest rate that is slightly higher than the one consistent with the risk profile of the talented innovators.[16] Given the terms of the optimal contract, the expected end-of-period wealth of a talented agent in this class is $\rho_H \left(w + RA\right)/\rho_L - e$. Finally, the unconstrained-class agents have $A \geq A_{HH}$. Only talented agents in this wealth class become innovators, with an expected income of $\rho_H \varpi - R(I - A) - e$, whereas untalented agents become workers.

Denote by $U_i(A, \Omega)$ the expected income level achieved by an $i$-type with wealth $A$ in an economy with state $\Omega$. From the end-of-period wealth of the agents in the static partial equilibrium, this lemma follows.

**Lemma 1.** *Let $\Delta(A, \Omega) \equiv U_H(A, \Omega) - U_L(A, \Omega)$. Then: (i) $\Delta(\cdot) \geq 0$, $\forall A$. Moreover, (ii) $\Delta'_A(\cdot) = 0$, $\forall A < A_e$; (iii) $\Delta'_A(\cdot) > 0$, $\forall A \geq A_e$; (iv) $\Delta(A, \Omega) \equiv \mathcal{V}^{FB}(A, \Omega) - \mathcal{W}^{FB}(A, \Omega)$, $\forall A \geq A_{HH}$. Furthermore, (v) $\Delta'_Q(\cdot) = 0$, $\forall A \in (0, A_e) \bigcup (A_L, A_{HH})$; (vi) $\Delta'_Q(\cdot) > 0$, $\forall A \in [A_e, A_L] \bigcup (A_{HH}, \infty)$.*

Lemma 1 shows that the talent premium is weakly increasing in wealth and machines' quality, and only reaches its full information counterpart in the unconstrained class. This means that talented agents benefit more from an increase in wealth and/or average quality than untalented agents.

**IV.2.2 Equilibrium Output and Growth.** Proposition 1 implies that the number of talented, $n_H$, and untalented innovators, $n_L$, in the static partial equilibrium are given by, respectively, $n_H(\Omega) = \lambda\left[1 - \Phi\left(A_e(\Omega)\right)\right]$ and $n_L(\Omega) = (1 - \lambda)\left[\Phi\left(A_L(\Omega)\right) - \Phi\left(A_e(\Omega)\right)\right]$. The total number of innovators is $n(\Omega) = n_H(\Omega) + n_L(\Omega)$, and thus the total number of workers is $1 - n(\Omega)$. Let $\rho \equiv (\rho_H n_H + \rho_L n_L)/n$ be the economy-wide average probability of success of the innovators. The expected number of successful innovations is given by $\rho n$: for each of these machines, initial quality improves by a factor $1 + \gamma$. Conversely, the expected number of unsuccessful innovations is $1 - \rho n$: for these machines, quality does not increase. As a consequence, in the static partial equilibrium, the growth rate of average quality is given by $g = \gamma \rho n$. The following lemma outlines some comparative statics of the static partial equilibrium.

**Lemma 2.** *(i) Consider two identical economies, but for the initial average quality of the machines, such that $Q_{-1} > Q'_{-1}$. Then $n_H \geq n'_H$ and $n_L \geq n'_L$, therefore $g \geq g'$*

[16] Precisely, banks ask for an interest rate of $\left(\rho_L \varpi - w - RA\right)/\left(\rho_L\left(I - A\right)\right)$ on these loans.

*and $Q \geq Q'$; (ii) Consider two identical economies but for the initial wealth distributions, $\Phi(A)$ and $\Phi'(A)$, such that $\Phi(A)$ first-order stochastically dominates $\Phi'(A)$. Then $n_H \geq n'_H$. (iii) Consider two identical economies but for the initial wealth distribution, such that $\Phi'(A)$ is a mean-preserving spread of $\Phi(A)$. Then if $\bar{A} < A_e$, $n_H \leq n'_H$; if $\bar{A} > A_L$, $n_H \geq n'_H$.*

The proof of Lemma 2 is straightforward. Part (i) says that, other things equal, more technologically advanced economies grow faster. This is because the reward to a successful innovation is greater, which makes becoming an innovator more attractive. Note that, whereas the wealth threshold $A_e$ is strictly decreasing in the profit of the successful innovator, $A_L$ (and $A_{HH}$) is strictly increasing in it: the more technologically advanced economy will have a larger lemon-class and a smaller working-class, and thus more innovators. Since the growth rate is increasing in the number of innovators, the more technologically advanced country grows faster, even if some of the additional innovators are of low ability.[17] Part (ii) says that, other things equal, wealthier economies tend to have more talented innovators. This is because, as the economy becomes wealthier, more agents will find themselves in the upper classes, where the adverse selection problem turns into an efficient redistribution (rich-class) or disappears (unconstrained-class). Part (iii) suggests that, when total wealth is low, it may be necessary to have an unequal distribution to allow talented agents to become innovators, otherwise all agents may find themselves in the working class; for wealthy economies, more equal wealth distributions are associated with more talented innovators.

## IV.3 A Sketch of Dynamics

The analysis in the previous section has been conducted within a static framework, as the quality of the machines and the wealth distribution at the beginning of the period were taken as given. Since these are actually endogenous, and reciprocally influence each other over time, here I present the dynamics of $Q_t$ and wealth.

Given the utility function in (1), individuals will optimally bequeath a fraction $\delta$ of their end-of-period income to their offspring. This amount will in turn fully determine the initial wealth of the new individuals. Henceforth, I split the population of agents

[17] Indeed, it is possible for the growth rate in the constrained equilibrium to be greater than the growth rate under full information. This happens if $\rho_H(\lambda - n_H) - \rho_L n_L$, a measure of the cost of having displaced the poor talented innovators, is negative. That having too many innovators may hurt the economy becomes clearer when we consider total output. Since the number of successful innovations corresponds to the number of monopolistically produced machines with demand given by (7), whereas the remaining machines have demand given by (9), the production of final good in the static partial equilibrium is equal to $y = l\,(\alpha p/\psi)^{\frac{\alpha}{1-\alpha}}\,\hat{Q}$, where $\hat{Q} \equiv \left(\rho n\left(\alpha^{\frac{\alpha}{1-\alpha}} - 1\right) + 1\right) Q$ is the average corrected quality of the machines at the end of the period, which takes into account that certain machines are produced competitively and others monopolistically. Whereas $Q$ is monotonically increasing in $\rho n$, $\hat{Q}$ is hump-shaped.

in lineages indexed by $h \in [0,1]$. Since types are intergenerationally uncorrelated by assumption, the wealth transition equations for any lineage are given by

$$A_{t+1,h} = \delta\left[w_t + RA_{t,h}\right] \qquad \text{if } A_{t,h} < A_{t,e};$$

$$A_{t+1,h} = \begin{cases} \delta\left[\varpi(\Omega_t) - R_t(I_t - A_{t,h})/\bar{\rho} - e\right], & \lambda\rho_H \\ 0, & 1-\bar{\rho} \\ \delta\left[\varpi(\Omega_t) - R_t(I_t - A_{t,h})/\bar{\rho}\right], & (1-\lambda)\rho_L \end{cases} \qquad \text{if } A_{t,h} \in \left[A_{t,e}, A_{t,L}\right];$$

$$A_{t+1,h} = \begin{cases} \delta\left[\left(w_t + R_t A_{t,h}\right)/\rho_L - e\right], & \lambda\rho_H \\ 0, & \lambda(1-\rho_H) \\ \delta\left[w_t + R_t A_{t,h}\right] & (1-\lambda) \end{cases} \qquad \text{if } A_{t,h} \in \left[A_{t,L}, A_{t,HH}\right];$$

$$A_{t+1,h} = \begin{cases} \delta\left[\varpi(\Omega_t) - R_t(I_t - A_{t,h}) - e\right], & \lambda\rho_H \\ 0, & \lambda(1-\rho_H) \\ \delta\left[w_t + R_t A_{t,h}\right] & (1-\lambda) \end{cases} \qquad \text{if } A_{t,h} \in \left[A_{t,HH}, I_t\right];$$

$$A_{t+1,h} = \begin{cases} \delta\left[\varpi(\Omega_t) + R_t(A_{t,h} - I_t) - e\right], & \lambda\rho_H \\ \delta\left[R_t(A_{t,h} - I_t)\right], & \lambda(1-\rho_H) \\ \delta\left[w_t + R_t A_{t,h}\right] & (1-\lambda) \end{cases} \qquad \text{if } A_{t,h} \geq I_t.$$

The dynamic path of the economy is dictated by the following system:

$$Q_t = (1 + \gamma\rho_{t-1}n_{t-1})Q_{t-1} \tag{18a}$$

$$\Phi_t(A) = \Gamma_{t-1}\left[\Phi_{t-1}(A)\right], \tag{18b}$$

where the operator $\Gamma_{t-1}[\cdot]$ maps the wealth distribution in $t-1$ into the initial wealth distribution in $t$, given the transition equations above. This operator evolves over time, as the transition equations depend on the average quality of the machines. The dynamic evolution of $Q_t$, in turn, depends on the wealth distribution, through the occupational choices of the agents. As a consequence, the dynamic system in (18) is non-stationary, and thus complicated to study. However, I can still prove the following results.

**Lemma 3.** *Consider two economies with identical initial wealth distribution,* $\Phi_t(A) = \Phi'_t(A)$. *Suppose also that* $Q_t > Q'_t$. *Then,* $\Phi_{t+1}(A)$ *first-order stochastically dominates* $\Phi'_{t+1}(A)$.

## V. The Empirical Analysis

In this section, I analyse the relationship between inequality of opportunity, inequality of effort, and innovation at the US state level. I first calculate various measures of inequality using the Panel Study of Income Dynamics (PSID). I then measure innovation in each state using data from the United States Patent and Trademark Office (USPTO). Finally, I empirically characterize the effect of inequality on innovation.[18]

## V.1 Inequality Measures

It is well-known that measuring inequality is empirically challenging (e.g. Keeley, 2015). Moreover, inequality measures are seldom comparable across countries. Data on wealth inequality, especially, is hard to come by; when available, it does not go back in time very far. On the contrary, data requirements to study the long-term effects of inequality of opportunity are very stringent: one not only needs comparable measures of inequality but also information for at least two distant periods in time, generally ten years (e.g. Marrero and Rodríguez, 2013). Given these limitations, I decide to carry our analysis at the US state level, using data from the PSID.

The PSID is the world's longest running household panel survey: it started in the 1968, with over 18,000 individuals living in 5,000 families in the United States, and it is still running. I use the weights supplied by the PSID to make the sample representative at the national level, by compensating for unequal selection probabilities and differential attrition.[19] Our analysis, however, is run at the state level. The drawbacks of using the PSID are that, unfortunately, the samples may not be representable at the state level, and that state sample sizes are small. To limit the impact of these problems, I drop states with fewer than 50 observations in a given year. This results in an unbalanced panel of 32 states distributed throughout the whole US territory: West (Arizona, California, Colorado, Oregon, Utah, Washington), South (Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, plus District of Columbia), Midwest (Iowa, Illinois, Indiana, Michigan, Minnesota, Missouri, Ohio, Wisconsin), and Northeast (Connecticut, Massachusetts, New Jersey, New York, Pennsylvania).

Since some information is not available for wives in all waves, I restrict our attention to individuals who are household heads (male in married family unit, but also female otherwise). I only consider individuals aged between 18 and 65 at the time of the interview. I calculate our measures of inequality based on the labour income of the respondents, given that information about other sources of income is not consistently available in the PSID. To account for composition effect, I first regress gross labour income on a second-order polynomial of potential experience. I then collect residuals from these regressions, and since they are centered around zero, I add a constant to match the minimum of the series.

Recycling notation from the theoretical model, let $x_i$ be the so-calculated gross income

[18] In this paper, I used Stata 15 by StataCorp (2017) and the user-written program by Liao (2016c).

[19] For longitudinal consistency, I disregard the Latino sample, that was added to the PSID data only between 1990 and 1995.

of individual $i = \{1, \ldots, N\}$ in a given state $s$ and year $t$ (for ease of reading, I drop these subscripts), $\bar{x}$ the weighted mean income of the state-year sample, and $f_i$ the sampling fraction of $i$ in the state-year sample (i.e. $i$'s sampling weight over the weight's sum for state $s$ in year $t$). For each state $s$ in year $t$, I estimate the corresponding Theil's $T$ index. This is defined as

$$T = \sum_{i=1}^{N} f_i \frac{x_i}{\bar{x}} \ln\left(\frac{x_i}{\bar{x}}\right). \tag{19}$$

Theil (1967) argued that this measure of entropy, or degree of disorder, provides a useful device for measuring inequality.[20] Theil's measure has been widely used in social science: one reason for its popularity is that, unlike the Gini coefficient, the total amount of inequality can be additively decomposed into a between-group component and a within-group component (see e.g. Liao, 2016a and 2016b). For this purpose, I partition the individuals in a given state and year into a mutually exclusive and exhaustive set of types, based on their father's education (i.e. no education, primary, secondary, and tertiary education) and race (i.e. white and non-white). I thus obtain (up to) eight types: all individuals in each type $m$ share the same *circumstances*. Race and parental education, as proxies for more general socio-economic background (e.g. wealth, transmission of ability and connections, investments in human capital) are the circumstances most widely used in the empirical literature. The between-group inequality component for a given state in a given year is calculated as

$$T_b = \sum_{m=1}^{M} y_m \ln \frac{\bar{x}_m}{\bar{x}}, \tag{20}$$

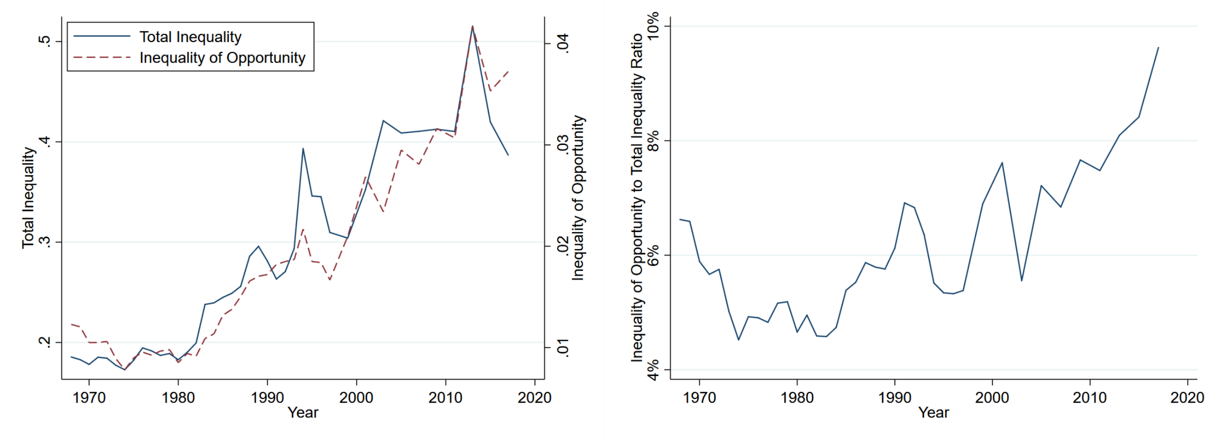
where $y_m$ is type $m$'s weighted income share expressed as a proportion of the weighted sample total income, and $\bar{x}_m$ is the weighted mean income of group $m$.

To summarise, $T_b$ measures inequality due to differences between circumstances: since these are beyond the individual's control, $T_b$ is used in the social sciences as a proxy for inequality of opportunity. Conversely, the within-group component, $T_w = T - T_b$, expresses inequality within types, and is thus seen as a proxy for inequality due to individual's choices or effort (over which the individual has control).

**V.1.1 Inequality in the US.** In this section, I briefly comment on these indexes. Figure 2a shows that both total inequality and inequality of opportunity at the US level were relatively stable up until the 1980s, while they have increased since then. The trend for total inequality is consistent with well-known facts, see e.g. US Census Bureau or Solt (2019). Figure 2b shows that the percentage of total inequality due to different

[20] Hereafter, Theil index refers to Theil's first measure, or Theil's $T$. There is actually a second measure, or Theil's $L$, which I do not employ. I estimate Theil's $T$ and its decomposition non-parametrically, like in Marrero and Rodríguez (2013), rather than parametrically, like in Ferreira and Gignoux (2011), given the structure of our database.

opportunities (arising from race and parental education) is modest but not insignificant (similarly to e.g. Ferreira and Gignoux, 2011 for six countries in Latin America, Marrero and Rodríguez, 2012, for 23 European countries, and Marrero and Rodríguez, 2013 for 26 US states), and seems to have been moving upwards in recent decades. Since one can realistically control for only a limited set of circumstances, $T_b$ is actually a lower-bound on the real inequality of opportunity (Ferreira and Gignoux, 2011, Marrero and Rodríguez, 2013).




(a) Values (b) Ratio

Figure 2: Time Evolution of Inequality, US

*Notes.* The total inequality index (Theil's $T$) and its decomposition are calculated by the author using data from the Panel Study of Income Dynamics.

## V.2 Innovation

Our measure of innovation builds on patent data. A patent is an exclusionary right conferred for a set period to the patent holder, in exchange for sharing the details of the invention. In the US, the USPTO is the agency that issues patents. Since 1976, it has provided information on the state of residence of the inventors, citation links between individual patents, and application year.

From the great amount of information available from the USPTO, Aghion *et al.* (2019, Supplementary Data) provide a ready-to-use dataset containing information on utility patents granted between 1976 and 2009 (up to 2006 when using quality-adjusted measures). In particular, for each state and year, they provide the flow numbers of patents, both as is and weighted for various proxies for a patent's quality, like the number of citations received.

The measure of innovation that I employ below is the number of patents granted, weighted by the number of citations received within 5 years of the application date, and corrected for the different propensity to cite in different sectors and time periods. A patent is associated with the state of residence of the patent inventor; a patent is split proportionally across states if co-inventors live in different states.[21]

## V.3 Results

I are now ready to look at the effect of inequality on innovation. Our estimated equation is

$$log\,(innov_{i,t}) = \beta_1 log\,(ineq_{i,t-10}) + \beta_2 \boldsymbol{x}_{i,t-10} + \alpha_i + \epsilon_{i,t}, \tag{21}$$

where $innov_{i,t}$ is the flow of (quality-adjusted) patents per capita in state $i$ in year $t$, $ineq_{i,t-10}$ is a vector of inequality indexes in year $t-10$, and $\boldsymbol{x}_{i,t-10}$ is a vector of control variables. The error term has two components: $\epsilon_{i,t}$ is an idiosyncratic error, whereas $\alpha_i$ captures unobservable heterogeneity across states that is invariant across times. The vector of controls is parsimonious and includes only the unemployment rate (to control for the business cycle), lagged GDP per capita (in logs), and the growth rate of total population: whereas the resulting estimation suffers from omitted-variable bias problems, I avoid introducing significant collinearity problems.

I first estimate equation (21) using Pooled OLS. I then implement a within regression, and thus estimate the relationship between the differential growth in innovation across states and the differential growth in inequality. Whereas OLS ignore the error structure, the fixed effect technique is problematic because it relies mostly on within-state variability.[22]

Results are presented in Table 1, where the first two columns refer to the OLS estimation, whereas the second two columns refer to the FE estimation. In columns 2 and 4, I break down total inequality into the between-group component (the inequality of opportunity term) and the within-group component (the inequality of effort term). By including the inequality of opportunity term, I control for the observed circumstances, i.e. father's education and race. I see that, whereas most terms are insignificant when I employ FE, in the OLS framework the coefficient of the between component is strongly significant and negative, whereas the within-group term is associated with a significantly positive effect. In particular, a one standard deviation increase in our inequality of opportunity measure is associated with a 22 point decrease in our measure of innovation; conversely, a one standard deviation increase in the within component is associated with a 20 point increase.[23] These effects are hidden behind a positive, but smaller in magnitude

[21] See Aghion *et al.* (2019, Section 3.1.2) for more information. Results are similar when I use only the number of patents and when I adjust for different quality-measures: these are available on request.

[22] Panizza (2002) suggests regressing inequality on time and state dummies, and use the resulting R-squared measure as a proxy of within-state and within-period variability. R-squareds close to 0.60 in our panel indicate that indeed our inequality measures mostly vary cross-sectionally.

[23] These results include only those states that have at least 50 observations when calculating the Theil's indexes, but the results are robust to selection criteria equal to 20, 30, or 100 observations. Moreover, signs are robust to changes in the lags of the regressors to 5, 15, or 20 years.

and less significant, coefficient when I consider only total inequality.

Table 1: Innovation and Inequality, Per Capita Number of Patents (Citations-Adjusted)

| | OLS | | FE | |
|---|---|---|---|---|
| Total Inequality | 0.71** | | 0.06 | |
| | (0.33) | | (0.12) | |
| Inequality of Opportunity | | -7.40*** | | 0.18 |
| | | (1.42) | | (0.30) |
| Inequality of Effort | | 2.71*** | | 0.02 |
| | | (0.38) | | (0.14) |
| GDP per capita | 0.96*** | 1.03*** | -0.07 | -0.06 |
| | (0.16) | (0.14) | (0.18) | (0.18) |
| Population Growth | 1.75*** | 1.98*** | -0.59* | -0.60* |
| | (0.21) | (0.22) | (0.32) | (0.32) |
| Unemployment | -0.11*** | -0.11*** | 0.05*** | 0.05*** |
| | (0.02) | (0.02) | (0.01) | (0.01) |

Cluster robust standard errors are reported in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A constant and time dummies are included

## VI. Conclusions

In this paper, I studied the relationship between innovation, inequality of opportunity, and inequality of effort. I have taken many simplifying assumptions to keep the model tractable, e.g. (i) I are assuming that abilities and wealth are uncorrelated, and thus that abilities are also intergenerationally uncorrelated; (ii) I have taken the riskless interest rate and the wage rate as exogenously given, which abstracts from important sources of general equilibrium effects (see e.g. Grüner, 2003, Inci, 2013, Spiganti, 2018); (iii) I have only two types of agents, whereas it is well-know that introducing more types may lead to different policy implications. Moreover, our empirical analysis lacks a number of important features: (i) I have not controlled for conditional convergence across states; (ii) our results suffer from omitted variable bias;[24] (iii) our standard errors are not robust to autocorrelation; (iv) endogeneity problems may be present. Future researchers could try to address these shortcomings.

[24]For example Akcigit *et al.* (2017) argue that population density and financial development are important determinants of innovation. As explained by Panizza (2002), one should compare the results obtained here with those resulting for a model with more covariates, which may suffer from collinearity problems.

## References


Acemoglu, D., P. Aghion, L. Bursztyn, and D. Hémous. (2012), The environment and directed technical change, *American Economic Review 102*, 131 – 166.

Acemoglu, D., J. A. Robinson, and T. Verdier. (2017), Asymmetric growth and institutions in an interdependent world, *Journal of Political Economy 125*, 1245 – 1305.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen. (2018), The social origins and IQ of inventors, Mimeo.

Aghion, P., and P. Howitt. (2009), *The Economics of Growth*, MIT Press, Cambridge.

Aghion, P., and J. Tirole. (1994), The management of innovation, *Quarterly Journal of Economics 109*, 1185 – 1209.

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hémous. (2019), Innovation and top income inequality, *Review of Economic Studies 86*, 1 – 45.

Aghion, P., and P. Howitt. (1992), A model of growth through creative destruction, *Econometrica 60*, 323 – 351.

Akcigit, U., J. Grigsby, and T. Nicholas. (2017), The rise of American ingenuity: Innovation and inventors of the golden age, Mimeo.

Andreoni, J. (1989), Giving with impure altruism: Applications to charity and Ricardian equivalence, *Journal of Political Economy 97*, 1447 – 1458.

Andreoni, J. (1990), Impure altruism and donations to public goods: A theory of warm-glow giving, *The Economic Journal 100*, 464 – 477.

Banerjee, A. V., and A. F. Newman. (1993), Occupational choice and the process of development, *Journal of Political Economy 101*, 274 – 298.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen. (forthcoming), Who becomes an inventor in America? The importance of exposure to innovation, *Quarterly Journal of Economics*.

Caselli, F., and N. Gennaioli. (2013), Dynastic management, *Economic Inquiry 51*, 971 – 996.

Celik, M. (2018), Does the cream always rise to the top? The misallocation of talent in innovation, Mimeo.

DeMeza, D., and D. Webb. (1987), Too much investment: A problem of asymmetric information, *Quaterly Journal of Economics 102*, 281 – 292.

Ferreira, F., and J. Gignoux. (2011), The measurement of inequality of opportunity: Theory and an application to Latin America, *Review of Income and Wealth 57*, 622 – 657.

Foellmi, R., and J. Zweimüller. (2006), Income distribution and demand-induced innovations, *Review of Economic Studies 73*, 941 – 960.

Foellmi, R., and J. Zweimüller. (2017), Is inequality harmful for innovation and growth? Price versus market size effects, *Journal of Evolutionary Economics 27*, 359 – 378.

Galor, O., and J. Zeira. (1993), Income distribution and macroeconomics, *Review of Economic Studies 60*, 35–52.

Ghatak, M., M. Morelli, and T. Sjöström. (2007), Entrepreneurial talent, occupational choice, and trickle up policies, *Journal of Economic Theory 137*, 27 – 48.

Grüner, H. P. (2003), Redistribution as a selection device, *Journal of Economic Theory 108*, 194 – 216.

Hatipoğlu, O. (2012), The relationship between inequality and innovative activity: A Schumpeterian theory and evidence from cross-country data, *Scottish Journal of Political Economy 59*, 224 – 248.

Hellwig, M. (1987), Some recent developments in the theory of competition in markets with adverse selection, *European Economic Review 31*, 319 – 325.

Holmström, B. (1989), Agency costs and innovation, *Journal of Economic Behavior and Organization 12*, 305 – 327.

Inci, E. (2013), Occupational choice and the quality of entrepreneurs, *Journal of Economic Behavior & Organization 92*, 1 – 21.

Jaimovich, E. (2011), Sectoral differentiation, allocation of talent, and financial development, *Journal of Development Economics 96*, 47 – 60.

Jones, C., and J. Kim. (forthcoming), A Schumpeterian model of top income inequality, *Journal of Political Economy.*

Kambourov, G., and I. Manovskii. (2009), Occupational mobility and wage inequality, *Review of Economic Studies 76*, 731 – 759.

Keeley, B. (2015), *Income Inequality: The Gap between Rich and Poor*, OECD Insights, OECD Publishing, Paris.

Liao, T. F. (2016a), Evaluating distributional differences in income inequality, *Socius 2*, 1 – 14.

Liao, T. F. (2016b), Evaluating distributional differences in income inequality with sampling weights, *Socius: Supplements 2*, 1 – 6.

Liao, T. F. (2016c), THEILDECO: Stata module to produce refined Theil index decomposition by group and quantile, Statistical Software Components, Boston College Department of Economics.

Lo Bello, S., and I. Morchio. (2016), Like father, like son: Occupational choice, intergenerational persistence and misallocation, Mimeo.

Manso, G. (2011), Motivating innovation, *The Journal of Finance 66*, 1823 – 1860.

Marrero, G. A., and J. G. Rodríguez. (2012), Inequality of opportunity in europe, *Review of Income and Wealth 58*, 597 – 621.

Marrero, G. A., and J. G. Rodríguez. (2013), Inequality of opportunity and growth, *Journal of Development Economics 104*, 107 – 122.

Murphy, K. M., A. Shleifer, and R. L. Vishny. (1989), Income distribution, market size, and industrialization, *Quarterly Journal of Economics 104*, 537 – 564.

Panizza, U. (2002), Income inequality and economic growth: Evidence from American data, *Journal of Economic Growth 7*, 25 – 41.

Rodríguez, J. (2008), Partial equality-of-opportunity orderings, *Social Choice and Welfare 31*, 435 – 456.

Rothschild, M., and J. Stiglitz. (1976), Equilibrium in competitive insurance markets: An essay on the economics of imperfect information, *Quarterly Journal of Economics 90*, 629 – 649.

Solt, F. (2019), Measuring income inequality across countries and over time: The standardized world income inequality database, Feb, OSF. https://osf.io/3djtq. SWIID Version 8.0, February 2019.

Spiganti, A. (2018), Wealth inequality, innovative talent, and occupational choices, Mimeo.

Spiganti, A. (forthcoming), Can starving start-ups beat fat labs? A bandit model of innovation with endogenous financing constraint, *Scandinavian Journal of Economics*.

StataCorp. (2017), Stata Statistical Software: Release 15, College Station, TX: StataCorp LP.

Theil, H. (1967), *Economics and Information Theory*, North Holland, Amsterdam.

Tirole, J. (1988), *The Theory of Industrial Organization*, MIT Press, Cambridge.

Violante, G. L. (2008), Skill-biased technical change, In *The New Palgrave Dictionary of Economics*. Eds. by S. Durlauf, and L. Blume.

Wilson, C. (1977), A model of insurance markets with incomplete information, *Journal of Economic Theory 16*, 167 – 207.

Zweimüller, J. (2000), Schumpeterian entrepreneurs meet Engel's law: The impact of inequality on innovation-driven growth, *Journal of Economic Growth 5*, 185 – 206.

# A. Appendix

## A.1 Glossary of variables and parameters

Variables and parameters definition

| | |
|---|---|
| $c_{t,i}$ | consumption of the final good |
| $b_{t,i}$ | bequest |
| $\delta$ | budget share |
| $\lambda_t$ | proportion of talented agent |
| $A$ | wealth endowment |
| $y_t$ | final good |
| $\alpha$ | output share |
| $l_t$ | labour |
| $x_{t,m}$ | quantity of machine $m$ |
| $Q_{t,m}$ | quality of machine $m$ |
| $p_t$ | price of the final good |
| $w_t$ | wage rate |
| $r_{t,m}$ | price of machine $m$ |
| $\rho_i$ | probability of innovating |
| $\gamma$ | increase in quality after innovation |
| $R_t$ | riskless rate of return |
| $I_t$ | setup cost |
| $D_i^s$ | repayment in realised state $s$ |
| $V_i^s$ | realised payoff of the innovator |

## A.2 Proofs and Maths

*Monopoly price is a constant markup over cost.* Define the point elasticity with respect to demand as

$$e = \frac{\partial Q}{\partial P} \times \frac{P}{Q} \rightarrow \frac{\partial Q}{\partial P} = e \times \frac{Q}{P},$$

where $Q$ is quantity and $P$ is price. Express the profit function of the monopolist as $\pi = PQ(P) - cost(Q(P))$ so that the FOC with respect to price is

$$\frac{|e|}{|e|-1} \times MC = P^{\star}.$$

Here, $MC = \psi$ and $e = -1/(1-\alpha)$ and thus $r_m = \psi/\alpha$. □

*Proof of Proposition 1.* Below, we derive the various contracts that banks offer to given wealth classes. The proofs are similar to any adverse selection model in financial markets (e.g. Grüner, 2003, Jaimovich, 2011, Inci, 2013).

For generality, we do not impose the repayment in the failure state equals to zero: nevertheless, given limited liability, this must be non-positive. A general contract then is

$$\boldsymbol{\sigma}_z(A,\Omega) = \begin{bmatrix} \sigma_H(A,\Omega) \\ \sigma_L(A,\Omega) \end{bmatrix} = \begin{bmatrix} D_H^S(A,\Omega) & D_H^F(A,\Omega) \\ D_L^S(A,\Omega) & D_L^F(A,\Omega) \end{bmatrix}, \tag{A.1}$$

where $D_i^S$ and $D_i^F$ are the repayments to the bank by the $i$-type agent in the success and failure state, respectively. Let the net payoff of an $i$-type innovator in the success state be given by $V_i^S = \varpi(\Omega) - D_i^S(A,\Omega) - e_i$; limited liability implies $V_i^S \geq 0$. Let the net payoff of an $i$-type innovator in the failure state be given by $V_i^F = -D_i^F(A,\Omega) - e_i$; limited liability implies $V_i^F \geq -e_i$. Thus, the expected payoff of an innovator, is

$$\mathcal{V}_i(A,\Omega) \equiv \rho_i V_i^S + (1-\rho_i)\, V_i^F, \quad \forall i = \{H, L\}.$$

The expected payoff of an agent who becomes a worker is $\mathcal{W}_i(A,\Omega) \equiv w + RA, \forall i = \{H, L\}$: we call this payoff "the outside option" to innovation.

The zero-profit conditions from separating contracts are

$$\rho_H \left(\varpi(\Omega) - V_H^S\right) - (1-\rho_H)\, V_H^F = R(I-A) \tag{A.2a}$$

$$\rho_L \left(\varpi(\Omega) - V_L^S\right) - (1-\rho_L)\, V_L^F = R(I-A), \tag{A.2b}$$

where the first line refers to talented agents, and the second one to untalented agents with wealth $A$. For given levels of $\mathcal{V}_H$ and $\mathcal{V}_L$, respectively, the corresponding iso-profit lines of the borrowers are

$$\bar{\mathcal{V}}_H = \rho_H V_H^S + (1-\rho_H)\, V_H^F \tag{A.3a}$$

$$\bar{\mathcal{V}}_L = \rho_L V_L^S + (1-\rho_L)\, V_L^F. \tag{A.3b}$$

The zero-profit condition from a pooling contract is given by

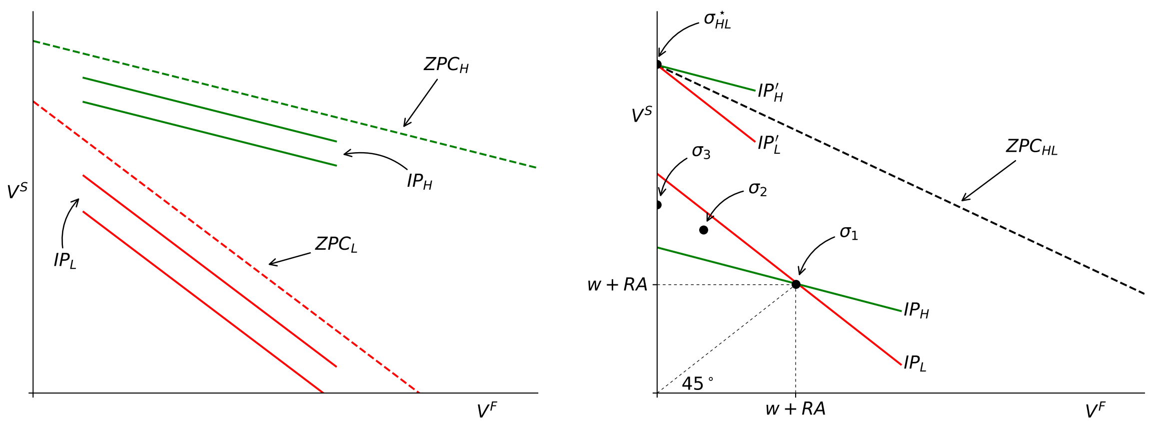
$$\rho_A D^S + (1-\rho_A)\, D^F = R(I-A), \tag{A.4}$$

where $D^S$ and $D^F$ are repayments of a random borrower with wealth $A$ in the success and failure state, respectively, and $\rho_A$ is her Bayesian probability of success.

The zero-profit conditions in (A.2) and groups of iso-payoffs in (A.3) are drawn in Figure A.1a. Each zero-profit condition $ZPC_i$ has the same slop as the corresponding iso-profit $IP_i$; agent's expected payoff is increasing as we move north-east, bank's profits are increasing as we move south-west. Imagine a bank offering two distinct contracts each on a zero-profit condition: the untalented agent would always pretend to be talented. Indeed, it is impossible to find a menu of contracts such that the zero-profit conditions hold and the untalented agents do not prefer the contract designed for the talented innovators.

Which contract can the banks then offer? It turns out that the equilibrium contract differs depending on the wealth class of the agents. Focus on Figure A.1b, where the outside option to innovation is given by the point $\boldsymbol{\sigma}_1$. The iso-profit curve for a talented and an untalented agent passing through this point are labelled $IP_H$ and $IP_L$, respectively. From the iso-profits in (A.3), we know that the iso-profit of the untalented agent is steeper. The banks could then offer any contract on the north-west of $\boldsymbol{\sigma}_1$, like $\boldsymbol{\sigma}_2$: under any reasonable belief, such a contract would attract only talented agents, and since we are below the zero-profit condition with only talented agents (not shown but it would be above $ZPC_{HL}$), the deviating bank would make positive profit. But any contract offered in this area can be undercut by another contract on its left. Due to limited liability, however, we cannot move further west than $V^F = 0$, like in $\boldsymbol{\sigma}_3$. But, similarly, $\boldsymbol{\sigma}_3$ can be undercut by any contract offering slightly better repayment in case of success. If such contract is below the zero-profit conditions with both types, $ZPC_{HL}$, a deviating bank would still make positive profits. Thus, undercutting goes on until banks make zero-profit, like in $\boldsymbol{\sigma}^\star_{HL}$, at which point no profitable deviation exists. Since this equilibrium pooling contract is on the vertical axis and lies on the zero-profit condition with both types given by (A.4), it follows that

$$\boldsymbol{\sigma}^\star_{HL}(A, \Omega) = \begin{bmatrix} R(I-A)/\rho_A & 0 \\ R(I-A)/\rho_A & 0 \end{bmatrix}. \tag{A.5}$$




(a) No Separating Contract (b) Pooling Contact

Figure A.1: Contracts

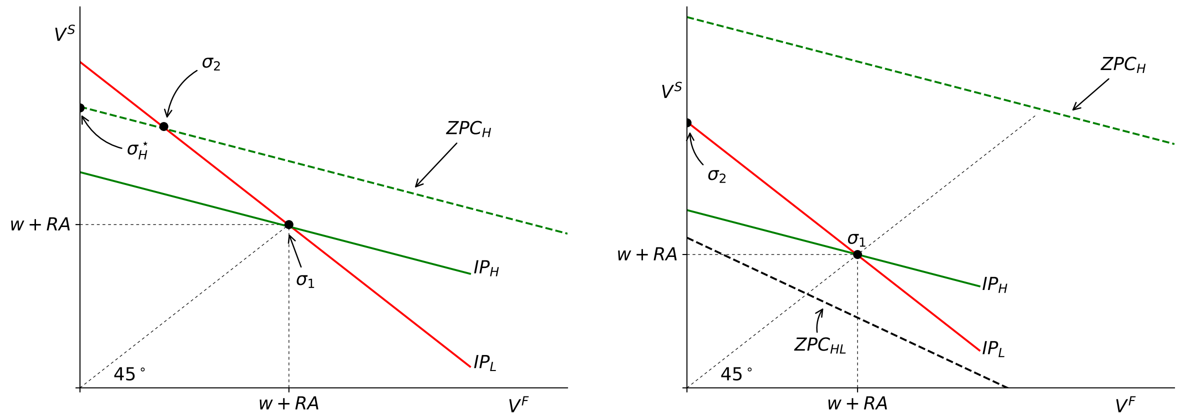
*Notes.* $IP_i$ are iso-payoffs, $ZPC_i$ are zero-profit conditions.

Now consider the situation illustrated in Figure A.2a. Agents are wealthy enough, so that their outside option is better than any pooling contract on the zero-profit condition with both types (not shown). A contract like $\boldsymbol{\sigma}_1$ cannot be an equilibrium: a deviating bank could offer any contract in the area between $IP_L$, $IP_H$, and $ZPC_H$: such contract would be accepted by talented agents only, and would thus entail positive profits. The equilibrium contract must thus lie on $ZPC_H$, whose equation is given in (A.2a). However, this time we have a continuum of equilibria in between $[\boldsymbol{\sigma}_2, \boldsymbol{\sigma}_H^\star]$. For simplicity, we choose to focus on the contract on the vertical axis,

$$\boldsymbol{\sigma}_H^\star(A, \Omega) = \begin{bmatrix} R(I-A)/\rho_H & 0 \\ R(I-A)/\rho_H & 0 \end{bmatrix}, \tag{A.6}$$

but this is without loss of generality since all these contracts entail the same expected payment and the same occupational choices.

Consider now Figure A.2b, which represents a wealth class for which banks can offer neither the zero-profit pooling contract $\boldsymbol{\sigma}_{HL}^\star$ nor the separating contract $\boldsymbol{\sigma}_H^\star$. The pooling contract on the zero-profit condition cannot be offered because it yields an expected payoff that is always lower than the outside option. A separating contract on $ZPC_H$ cannot be offered because it would also be accepted by untalented agents. Can the separating menu of contract $[\sigma_2, \sigma_1]^T$ be offered? Untalented agents are indifferent between $\sigma_1$, $\sigma_2$, and their outside option, so, by assumption, they choose to stay out of the innovation sector. Talented agents strictly prefer $\sigma_2$ to $\sigma_1$ and the outside option, and thus accept the contract. If banks can offer only loan contracts, then this menu of contract is an equilibrium in which banks make positive profits, since $\sigma_2$ is below the zero-profit condition with only talented agents, and untalented agents do not apply for loans.[25] This is the equilibrium contract we consider in the main text.




(a) Separating Contract (b) Profitable Separating Contract

Figure A.2: Separating Contracts

*Notes.* $IP_i$ are iso-payoffs, $ZPC_i$ are zero-profit conditions.

Assume for a moment that banks are not limited to loan contracts alone. Then banks can undercut each other by offering a contract to the talented agent that is slightly above $\sigma_2$ but still below $ZPC_H$, and by paying lenders in both states of the world something more than the usual interest on deposits. Undercutting goes on until banks make zero-profit on these contracts, like $[\sigma_3, \sigma_4]^T$ in Figure A.3. Since banks make profits on $\sigma_3$ and

[25] Indeed, there is no profitable deviation. A deviation contract below $IP_L$ is not accepted by anyone; one above $IP_L$ but below $IP_H'$ is only accepted by untalented agents; a contract with $V^F < 0$ would violate limited liability; any pair of contracts above $IP_H'$ would be accepted by everyone but would incur losses because it would be above the zero-profit condition with both types.

losses on $\sigma_4$, a Nash player would cancel $\sigma_4$: since the other banks are still offering it, though the deviating bank would be better off. A solution to this non-existence problem, is to impose a Wilson's (1977) equilibrium concept, where players are non-myopic rational. In a Wilson's (1977) world, the deviating bank would take into account the effects of its action on the actions of other banks. The non-myopic player knows that other banks would react to the cancelling of $\sigma_4$ by withdrawing $\sigma_4$ as well, and thus would incur losses: as a consequence, it would not deviate in the first place.[26]

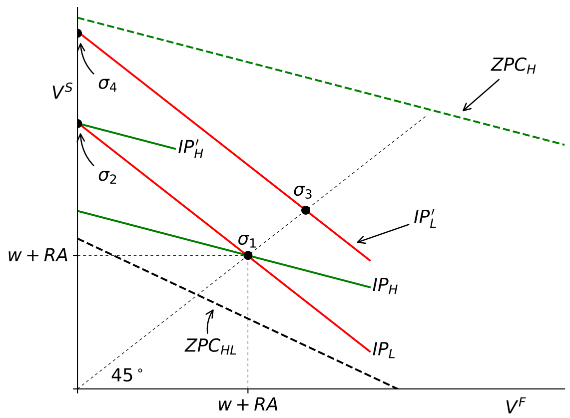



Figure A.3: Wilson's (1977) Separating Contract

*Notes.* $IP_i$ are iso-payoffs, $ZPC_i$ are zero-profit conditions.

Given the contracts derived above, and the participation and incentive compatibility constraints given in the main text, occupational choices are trivial. □

*Proof of Lemma 1.* Notice that

$$\Delta(A, \Omega) = \begin{cases} 0, & A \leq A_e \\ (\rho_H - \rho_L)(\varpi - R(I - A)/\bar{\rho}) - e, & A \in [A_e, A_L] \\ (w + RA)(\rho_H - \rho_L)/\rho_L - e, & A \in [A_L, A_{HH}] \\ \rho_H \varpi - R(I - 2A) - w - e, & A \geq A_{HH} \end{cases} \tag{A.7}$$

By visual inspection of (A.7), the talent premium is (i) strictly decreasing in $e$ in all classes but the working-class (where it is independent of $e$); (ii) unaffected by $w$ for the working and lemons-class, strictly increasing in $w$ for the riches, and strictly decreasing in $w$ for the unconstrained-class; (iii) independent of $R$ for the working-class, strictly decreasing in $R$ for the lemons-class, strictly increasing in $R$ for the rich-class, and could be non-monotonic for the unconstrained agents (it is strictly increasing for those agents who do not have to borrow from banks, and for the unconstrained borrowers only if $A_{HH} > 0.5I$); (iv) weakly increasing in the average quality of the machines at the beginning of the period (since the expected profits from innovations are increasing in the quality of the machines); (v) weakly increasing in $A$ both within and across classes (note that the talented premium is continuous across classes, and weakly increasing within classes). □

[26] Whether I impose a Nash-Bertrand equilibrium concept with a restricted set of contracts, or this Wilson's (1977) equilibrium concept with a larger set of feasible contracts, the qualitative results are unchanged.