

# Misallocation in the Market for Inventors\*

Fil Babalievsky

December 2, 2022

[Click here for latest version](#)

## Abstract

How much misallocation is there in the market for inventors? This paper introduces a novel growth model where firms search for inventors in a frictional labor market, and where the efficiency of innovation depends on the quality of the match between the inventor and firm. The model features both knowledge spillovers and congestion in the labor market, hence the resources invested in search can be either too low or too high. I use the model to quantify this tradeoff and to study what policies can rectify the inefficiencies in this market. I document that inventors are 40% more productive when moving across firms, a finding not driven by differences in firm innovativeness. I find that search frictions and the resulting misallocation of inventors across firms are of first-order importance for growth: shutting down search frictions and moving to a world where inventors are perfectly allocated across firms increases growth from 2.0 to 2.6 percent. Moreover, I find that firms invest too little in searching for inventors, and that conventional R&D subsidies do not address the inefficiencies in the market for inventors.

---

\*University of Minnesota. I thank Kyle Herkenhoff, Loukas Karabarbounis, Jeremy Lise, Lee Ohanian, and Michael Peters for their support and feedback. I thank Jacob Adenbaum, Cristián Aguilera Arellano, Mariacristina De Nardi, Nik Engbom, Martín García-Vázquez, Eugenia Gonzalez-Aguado, William Jungerman, Hannes Malmberg, Thomas May, Marta Morazzoni, Jo Mullins, Jane Olmstead-Rumsey, Jordan Pandolfo, Luis Pérez, Marta Prato, Nicolò Russo, Kjetil Storesletten, and the participants of the De Nardi-Lise-Mullins and Herkenhoff-Karabarbounis-McGrattan workshops at the University of Minnesota for many helpful conversations. All errors are my own.

## Introduction

This paper studies the labor market for inventors. Which inventor works for which firm can affect both firm performance and economic growth, as innovative activity combines inventor skills with the intangible capital of their firm. Hence, it does not come as a surprise that many high-profile inventors make headlines when they moved across firms (among them Peter Rawlinson from Tesla to Lucid and Jeffrey Wilcox from Intel to Apple and back)<sup>1</sup>. In this paper I propose a novel growth model that formalizes these stories and shows that the allocation of inventors has first-order consequences for firm dynamics and growth.

I make three contributions. The first is theoretical: my model takes seriously the idea that firms need to look for talented inventors before coming up with new products. In the model, firms hire workers ("headhunters") to search for inventors to match with their existing product lines. They then furnish the inventors with costly resources to innovate and help the firm expand into new product lines, creatively destroying the incumbent product line. If an inventor was working on a product line that was creatively destroyed, their match breaks. A key feature of the model is that inventor-firm matches are differentiated and may be of high or low quality. As both matched inventors and firms can search for replacements for their incumbent matches, both move up the "match quality ladder" over time until their matches are destroyed. This model builds on both the growth and search literatures, and it inherits some of the efficiency properties of both. As in all models of growth, firms benefit from knowledge spillovers from their rivals, and therefore do not internalize the benefit of making better matches and thereby innovating more quickly. However, when firms decide how much effort to devote to searching for inventors, they do not internalize the benefits to the inventors of being in a better match or the costs to rival firms of losing their inventors. Search frictions can, as usual, mean that firms expend either too little or too much effort to meet with workers. On the one hand, inventor mobility suggests greater returns to innovative effort at their new firms than their old, and these returns are not internalized by the new firm as it chooses how hard to search. On the other hand, the resources that their old firms invested<sup>2</sup> to recruit them in the first place went to waste—as usual in search models, the privately optimal level

---

<sup>1</sup>Details about Rawlinson and Wilcox's stories can be found [here](#) and [here](#), respectively.

<sup>2</sup>[Gavazza, Mongey, and Violante \(2018\)](#) provides a useful decomposition of recruiting costs, and finds that the two that comprise the largest share of such costs are payments to job boards or search engine aggregators (for example, purchasing a license to Monster's database) and payments to agencies or third-party recruiters.

of search effort may be greater or lower than the social optimum. These considerations are missing in standard models of growth, where the inputs into innovation are instead undifferentiated and hired in a frictionless market. Another novel feature of this model is that creative destruction does not simply transfer control of a product line from one firm to another, but also destroys the incumbent firm's match with its inventor. As products become obsolete, the inventors who were working on those product lines find that their skills are no longer valuable. This is a realistic negative consequence of creative destruction that is missing from most growth models: as firms innovate and render their rivals' products obsolete, they also diminish the innovative capacity of the economy by breaking their matches. This is the "Blackberry effect" —when Blackberry's smartphones were creatively destroyed by Apple's iPhone, many of its engineers found themselves unemployed.<sup>3</sup>

The key novel parameters in my model are those that govern the degree of search frictions and the costs of inventor misallocation (or, more formally, the efficiency of the meeting technology and the difference between the best and worst possible match qualities). My second contribution is therefore to use disambiguated patent data from the USPTO to document novel facts about inventor mobility across firms to discipline these parameters. In particular, I document that inventors move frequently across firms, and that these moves are associated with large increases in patenting. Taken together, these moments suggest a meaningful role for search frictions and misallocation. More specifically, I document that inventors work at roughly 1.6 firms on average, which suggests a significant amount of turnover given that the average time between an inventor's first and last patent is less than six years. The rate of mobility across firms helps discipline the degree of search frictions in the market for inventors. To discipline the costs of misallocation, I use a matched-pairs design related to that in [Prato \(2022\)](#). I compare inventors who move across firms at  $t = 0$  to a "twin" inventor who never moves but has the same first year in the patent data, had the same cumulative number of patents at  $t = -1$ , was working same firm as of  $t = -1$ , and had the same number of new patents at exactly  $t = -1$ . I find that the "treated" inventor who moves has about 40% more patents going forward. The fact that inventors see their productivity rise upon moving suggests a role for inventor-firm match quality, and the fact that mobility is fairly frequent among inventors suggests that inventors are not perfectly allocated across firms and lends credibility to a

---

<sup>3</sup>[This](#) news article describes some of their stories.

model with search frictions. This is not driven by sorting to firms that have systematically more patents per inventor, which also supports a role for inventor-firm specific match quality: if some firms truly were systematically better at taking the same inventors and getting more innovations out of them, one might expect that the increase in patenting upon a move would be mostly driven by the firm effect. This matched-pairs comparison provides striking suggestive evidence that mobility across firms is associated with more patenting, but it has no causal or structural interpretation in and of itself. Both the treatment and placebo are selected conditional on continuing to patent, which may introduce subtle selection effects. More importantly, I do not know if the treatment moved straight from job-to-job or if they transited through unemployment. Consider an inventor who climbed up the match quality ladder, lost their latest job, and then quickly matched to another one. Such an inventor would show up as a “treated” inventor in this analysis, and would bias down the estimates relative to the true average treatment effect of meeting and agreeing to move to a new employer. But I run the same regressions in both my model and the data, and use them as informative moments to discipline the underlying structural parameters.

My third contribution is quantitative: I show that my model has a rich set of implications for growth, policy, and firm dynamics. First, I use the estimated model to quantify the impact of the misallocation of inventors on growth. I find that removing this misallocation altogether could raise growth from 2.0 percent to 2.6 percent, which shows that search frictions in the market for inventors are costly for growth. Intuitively, the model finds such large costs because inventor mobility is both very frequent and also associated with an increase in patenting. I then show that introducing search frictions to a standard growth model in the vein of [Klette and Kortum \(2004\)](#) has significant implications for firm dynamics, and can help rationalize why some firms grow extremely quickly. [Jones and Kim \(2018\)](#) notes that workhorse random-growth models of firm dynamics do not generate firms like Microsoft or Google that grow to enormous sizes in a short span of time, and note that it is typically necessary to add some kind of heterogeneous growth rate across firms. They cite the “Luttmer Rocket” of [Luttmer \(2011\)](#), a short-lived period when certain firms have better ideas before reverting to the mean, as one way to match the speed at which superstar firms grow large. Introducing search frictions to the Schumpeterian growth framework also helps address this problem: in the calibrated model, most newborn firms never match to an inventor, and therefore never spin off any new product lines. The few firms that do match innovate and expand into new

product lines very quickly, which gives them the opportunity to match with more inventors down the line. Most firms in the calibrated model never get the chance to match with an inventor and never get to innovate, and the firms that do match need to innovate very quickly to rationalize the average rate of growth and creative destruction that we see in the data. Adding search frictions for inventors therefore helps generate *endogenously* heterogeneous growth rates, especially for newborn firms with relatively few product lines. The calibrated model can easily generate new firms like Amazon or Microsoft that become innovation leaders within one generation, in stark contrast to most firm-dynamics models. Notably, my model generates at least as much dispersion in firm innovation as seen in the data without resorting to any ex-ante heterogeneity among firms or inventors. The only dimension of heterogeneity is along inventor-firm match quality, and even this is limited.<sup>4</sup> There are no “superstars” or geniuses in my model—ex-post firm heterogeneity is purely a matter of luck.

Finally, I use the model to conduct a series of counterfactuals. First, I ask whether firms are investing too much or too little in search. To do so, I add distortions to the firms’ problem of how much search effort to engage in and re-solve the model. I find that consumption-equivalent welfare increases as search effort increases, meaning that firms underinvest in search. Two major factors driving this result are the fact that my calibrated search costs are very low, and the fact that there is decreasing returns to search at both the firm and aggregate level. Second, I test whether conventional R&D subsidies can address these inefficiencies by themselves. Intuitively, subsidizing research costs increases the net returns to hiring an inventor, and so one might hope that conventional research subsidies can address the fact that firms invest too little in search for inventors. In partial equilibrium this certainly holds to some extent, but in general equilibrium, it does not. Recall that all innovation in the model is creative destruction, and that an inventor working for a creatively-destroyed product line loses their job. Note that this mechanism is in addition to the effect of creative destruction on the net present value of profits earned from innovation, as detailed in [Peters \(2020\)](#). In that model as well as mine, a higher rate of creative destruction weakens firms’ incentives to invest in innovation and reduces the intensity of innovative effort. In my model, creative destruction also destroys some of the economy’s capacity to innovate. Greater subsidies for research give firms stronger incentives to search for inventors and spur

---

<sup>4</sup>The highest-quality match is only about 25 percent more innately productive than the worst.

innovation, but at the aggregate level an increase in innovation also means an increase in the rate at which “match capital” is destroyed. Remarkably, I find that these two forces cancel each other out almost perfectly in general equilibrium: as I raise the R&D subsidy to an unreasonably high share of GDP, both the share of matched inventors and the average match quality are basically unchanged. The partial equilibrium increase of a firm’s innovation in response to an R&D tax credit, as one might estimate in a quasi-experiment, is partly driven by improved match quality, but this mechanism does not scale up in general equilibrium. Modeling the process by which firms match to inventors is therefore important for translating partial-equilibrium estimates into general-equilibrium predictions about the impact of innovation policy.

## Literature and Contribution

This paper builds on and contributes to several literatures. First, this paper introduces a model of endogenous growth and firm dynamics in the tradition of [Klette and Kortum \(2004\)](#). A standard feature of such models is that firms can hire inventors (and other inputs into idea creation) on a frictionless spot market. My contribution is to instead model a frictional market for inventors and show that such a market structure has markedly different and empirically relevant consequences for policy and firm dynamics.

This paper also contributes to the literature on frictional search by applying it to inventors and disciplining it with patent data. Other papers bridging the search and Schumpeterian growth traditions are [Engbom \(2020\)](#) and [Lentz and Mortensen \(2012\)](#), but these papers model frictional search for production workers as opposed to inventors.

A growing line of research uses patent data to study inventors. [Prato \(2022\)](#) and [Arkolakis, Peters, and Lee \(2020\)](#) show that the migration of inventors across nations is a major factor in driving growth, whereas my paper studies the smaller-scale but more frequent mobility of inventors across firms. [Akcigit, Caicedo, Miguelez, Stantcheva, and Sterzi \(2018\)](#) studies the allocation of inventors across research teams in a model where ideas are sold on a competitive market. [van der Wouden and Rigby \(2020\)](#) studies inventor mobility from 1900 to 1975 and also finds that inventor mobility is associated with an increase in patenting. I perform a similar empirical exercise and use it to discipline a structural model and perform policy counterfactuals. [Pearce \(2019\)](#) also finds an

important role for inventor-firm match quality in the rate of innovation.<sup>5</sup>

Many papers studying the determinants of firm growth have shown that

Finally, this paper contributes to the literature on innovation policy, in the vein of [Atkeson and Burstein \(2019\)](#). Another related and complementary paper is [Acemoglu, Akcigit, Alp, Bloom, and Kerr \(2018\)](#), which studies R&D policy in a model that has heterogeneous and dynamic firm types but frictionless markets for homogeneous skilled labor. In both that paper and mine, the key source of inefficiency is the presence of a knowledge spillover: holding all else equal, raising productivity at other firms increases the returns to innovation at one’s own firm. In both my paper and [Acemoglu et al. \(2018\)](#), growth depends on how the inputs into allocation (inventors or skilled labor) are allocated across firms, and in both papers the equilibrium allocation is generally not optimal. [Celik \(2022\)](#) considers a complementary growth model where not everyone can be an inventor due to limited training, and finds that individuals are misallocated across occupations.

## 1 Model

Time is continuous. There is a stand-in household with a measure  $M$  of workers, who inelastically supply their labor to profit-maximizing firms. An exogenous share  $\nu$  of these individuals work in the frictional market for inventors<sup>6</sup>, and the remaining share  $1 - \nu$  can freely move across four frictionless sectors: final goods production, new firm creation, “headhunting” for inventors, and creating the “research good”<sup>7</sup>. Inventors become non-inventors at a rate  $\delta_i$ , and non-inventors become inventors at a rate  $\delta_r = (\nu\delta_i)/(1 - \nu)$ .<sup>8</sup>

There is perfect risk-sharing in the household. The household derives logarithmic utility from the final good  $Y$  and discounts the future at rate  $\rho$ . Consumption of the final good is the only source of utility in this model and there are no aggregate shocks, hence we can write the household’s

---

<sup>5</sup>[Pearce \(2019\)](#) also notes a separate major role for the match quality between a new inventor and the firm’s existing team of inventors. [Prato \(2022\)](#) also studies the dynamics of teams of inventors. I do not focus on teams for the sake of tractability but view merging these strands of the literature as an important direction for future research.

<sup>6</sup>I denote this share  $\nu$  of individuals as “inventors” regardless of whether they are currently matched or not. The assumption of an exogenous measure of inventors is a reasonable simplification in light of the findings in [Bell, Chetty, Jaravel, Petkova, and Van Reenen \(2019\)](#) that the children of inventors are far more likely to be inventors themselves. See [Celik \(2022\)](#) for a model that endogenously generates a fixed share of inventors.

<sup>7</sup>This is a stand-in for all non-labor inputs into innovation

<sup>8</sup>This is consistent with a steady-state share  $\nu$  of inventors and helps to rationalize the fact that most inventors only patent for a few years.

discounted utility  $U$  as:

$$U = \int_{t=0}^{\infty} e^{-\rho t} \log(Y_t) dt.$$

## 1.1 Production

The production side of the model is deliberately standard. As in [Klette and Kortum \(2004\)](#), the final good is a CES aggregate of an exogenous measure 1 of varieties. Each variety  $i$  may be produced by multiple firms  $f$ , and each firm may produce more than one good. Different varieties  $i' \neq i$  are imperfect substitutes, with elasticity of substitution  $\sigma$ , but the same variety produced by different firms  $f' \neq f$  are perfect substitutes. I find it convenient to denote  $[fi]$  as the “division” within firm  $f$  that produces good  $i$ . Production of the final consumption good  $Y_t$  is as follows, where  $y_{[fi]t}$  is the production of each division:

$$Y_t = \left( \int_0^1 \left( \sum_f y_{[fi]t} \right)^{\frac{\sigma}{\sigma-1}} di \right)^{\frac{\sigma-1}{\sigma}}.$$

Division  $[fi]$  has access to a linear labor-only technology with productivity  $z_{[fi]t}$ :

$$y_{[fi]t} = z_{[fi]t} l_{[fi]t}.$$

I follow [Garcia-Macia, Hsieh, and Klenow \(2019\)](#) and assume that each division must pay an infinitesimal operating cost of  $\epsilon$  to produce. This ensures that only the division with the highest  $z$  in each product line  $i$  will produce, and that it will charge a homogeneous markup  $\mu = \sigma / (\sigma - 1)$ .<sup>9</sup>

I denote  $z_{it} = \max_f z_{[fi]t}$  as the efficiency of the leading producer of product line  $i$ ,  $F_t(z_{it})$  as the CDF of product lines with efficiency  $z$  at time  $t$ ,  $Z_t = \left( \int_0^1 z_{it}^{\sigma-1} dF_t(z_{it}) \right)^{\frac{1}{\sigma-1}}$  as the usual measure of “average productivity”, and the “relative efficiency” of a division as  $\hat{z} \equiv (z/Z)^{\sigma-1}$ .

I denote the total mass of workers involved in production as  $l_t$ . Final goods output  $Y$  is  $Z_t l_t$ , wages for production workers  $w$  are  $Y_t / (\mu l_t)$ , and product-level gross profits  $\pi$  can be written as:

---

<sup>9</sup>If two “divisions” belonging to different firms  $f$  and  $f'$  can produce the same good  $i$ , and both decide to pay the operating cost, they compete in Bertrand fashion. If  $z_{[f'i]t} < z_{[fi]t}$ , then division  $[f'i]$  will make no profits from production. Because of the operating cost, the less-productive division will not enter, and the more-productive division will charge a fixed markup.



$$\pi_{[fi]t} = (\mu - 1) \underbrace{\left( \frac{z_{[fi]t}}{Z_t} \right)^{\sigma-1}}_{\text{Relative Productivity } \hat{z}} w_t l_t.$$

In what follows, I drop time subscripts where there is no possibility of confusion.

## 1.2 New Firm Creation

Aside from creative destruction from inventors at incumbent firms, which is the main focus of this paper, growth can also come from creative destruction by entrants. There is free entry into new firm creation: any of the  $M(1 - \nu)$  non-inventors can at any time decide to try to start up a new firm, thereby converting their endowment of time into a flow rate of new firms.<sup>10</sup> The rate  $\phi_s$  at which a worker can start up firms depends on the quality of the entry technology  $\iota$  and the degree of congestion  $\psi > 0$  in entry: if a measure  $E$  of new firms are created, the productivity of the entry technology scales down by  $E^\psi$ , similar to [Peters and Zilibotti \(2021\)](#).

$$\phi_s = \iota / E^\psi.$$

Upon successfully entering, the entrepreneur<sup>11</sup> draws a random product line  $i$  and improves on the relative productivity  $\hat{z}_{[fi]t}$  of the leading producer of  $i$  by an additive factor  $\lambda$ . The entrepreneur creates a new firm  $f'$  with a single unmatched division to produce  $i$ . I assume that the “creatively destroyed” division  $[fi]$  exits and lays off their inventor.<sup>12</sup> The representative household’s valuation of the new startup is the net present value of its future profits.

Because innovations are cumulative, entrants do not internalize the impact of their efforts on future innovators. This means that there is an externality and that the market allocation will generally be inefficient.

I denote  $N(\hat{z})$  as the value of an unmatched division with relative productivity  $\hat{z}$ , and  $dF(\hat{z})$  as the PDF of product lines across relative productivity space. I can now write  $CD$ , which is both the

---

<sup>10</sup>Making the cost of entry be in units of labor rather than in units of the final good is important for achieving balanced growth.

<sup>11</sup>Note that entry is a separate technology from innovation in this model: entrepreneurs are drawn from the mass  $M(1 - \nu)$  of non-inventors, not the mass  $M\nu$  of inventors.

<sup>12</sup>This simplification greatly increases the tractability of the model and can be microfounded if I assume that the frontier division, armed with the knowledge of how to produce  $i$ , can copy any innovation made by the laggard, reducing their expected profits from innovation to zero.

expected value of creatively destroying a product line and the expected value of entry conditional on success:

$$CD = \int N(\hat{z}' + \lambda) dF(\hat{z}').$$

Because there are no frictions in or barriers between the labor market for final goods producers and the entrepreneurial market, entry  $E$  must be such that the returns to entry exactly balance out the opportunity cost of working in final goods production:

$$\phi_s CD = w.$$

The mass of workers  $e$  in entry must also be consistent with the level of entry  $E$ :

$$\phi_s e = E.$$

### 1.3 Innovation

Firms can only innovate if they match a division to an inventor. For tractability, I assume that each division within a firm can have at most one inventor.

I assume that inventors and research spending cannot be transferred across divisions within the same firm, and that search effort is division-specific. Combined with the fact that each variety is infinitesimal, this means that each division within a firm can make decisions without reference to any other division and that firm values are additively separable in division values.

All innovation in this model is creative destruction, which works the same for incumbents as for entrants: when the inventor and division innovate, they draw a random product  $i'$ , improve on its relative productivity by the same step size  $\lambda$ , and spin off a new unmatched division within the same firm to produce the improved product line. The expected value of an innovation is the same  $CD$  for an incumbent as it is for an entrant.

The rate of innovation is an increasing function of the resources  $r$  (in units of labor) which the division provides to the inventor. Research effort  $r$  is a stand-in for all kinds of R&D expenditures (lab technicians, equipment, etc). The rate of innovation is also increasing in the time-invariant division-inventor match quality  $q_{[fi]j}$ , which acts as TFP for the innovation process. This notion of

match quality is new to the literature.

I parameterize the rate of incumbent innovation  $\phi_n$  very simply<sup>13</sup>:

$$\phi_n(q, r) = \underbrace{q}_{\text{Match Quality}} \cdot \underbrace{r^\gamma}_{\text{R\&D}}.$$

Note that  $\hat{z}$  does not show up in the value of innovation, hence the benefit of matching to an inventor is constant across all values of  $\hat{z}$ .

## 1.4 Compensating Inventors and Investing in Research

To model the bargaining between the division and the worker, I assume the inventor and division can successfully bring the idea to market with probability 1, or the inventor can exercise their outside option and successfully bring the idea to market with probability  $\chi < 1$ . This captures the possibility that the inventor may try to leave the firm and steal the idea. If the division offers the inventor a share  $\chi + \epsilon$  of the value of the patent, the inventor will always give the patent to the firm and it will always be brought to market. I refer to  $\chi$  as the “bargaining weight” of the inventor.<sup>14</sup>

Divisions therefore anticipate that inventors will keep some share of the value of innovation, which induces a hold-up problem for the division. It must pay  $w$  for each unit of  $r$  to compensate those workers for their opportunity cost. The division therefore chooses  $r$  to maximize the following:

$$\max_r (1 - \chi)qr^\gamma CD - wr$$

Note that divisions do not compensate inventors with a flat wage, but rather only reimburse them upon a successful innovation occurring.<sup>15</sup> They cannot pre-commit to a piece rate  $w_i$  and will therefore never pay inventors more than a share  $\chi$  (as they would profitably deviate by cutting the

<sup>13</sup>Standard models in the tradition of [Klette and Kortum \(2004\)](#), such as [Peters \(2020\)](#), also assume a firm-level innovation production function that features constant returns to scale in *both* resources  $r$  and the number of product lines controlled by the firm.

<sup>14</sup>This is isomorphic to imposing an exogenous piece rate of  $\chi$ . I do not use a more-standard Nash bargaining setup as it generally cannot be solved as a convex optimization problem in a model with on-the-job search—see [Shimer \(2006\)](#). The key difference between this assumption and a typical Nash bargaining setup with bargaining parameter  $\chi$  is that it breaks the link between compensation and the expected duration of the match, which is the major source of difficulty in using Nash with on-the-job search. My approach still avoids the need to carry around an extra state, and as I will show later, it also helps avoid the need to solve for the inventors’ value functions altogether. Hence, it is more tractable than [Postel-Vinay and Robin \(2002\)](#).

<sup>15</sup>[Celik and Tian \(2022\)](#) and [Koh, Santaaulàlia-Llopis, and Zheng \(2020\)](#) document that inventors face high-powered incentives in the real world, which makes this assumption more defensible.

payment) or less than  $\chi$  (as the inventor would not share the idea with them.)

The aggregate rate of innovation per product line is equal to the rate of creative destruction, summed across entrant and incumbent innovation. I denote it as  $\Delta$ :

$$\Delta = \phi_s e + \int_{\hat{z}, q} \phi(q) dF^V(\hat{z}, q).$$

## 1.5 Matching to Inventors

I now explain how inventors are matched to divisions. To look for inventors, divisions hire headhunters  $h$  from the mass of non-inventors, paying them a wage  $w$  to bid them away from production, entry, and research. Effective search effort is  $m(h) = h^\eta$ , and if  $\eta < 1$  this implies diminishing returns to search at the level of the division. Divisions can search regardless of whether they are already matched or not; i.e. they have the option to fire their current inventor and replace them.<sup>16</sup> Search is random, and matched inventors are just as likely to be contacted as unmatched. Unmatched inventors do not earn unemployment insurance.

There is a constant-returns-to-scale aggregate meeting function that converts search effort (the number of unmatched divisions weighted by their search effort  $h^\eta$ ) to meetings  $H$ . This meeting technology is a Cobb-Douglas function of aggregate search effort  $S$  and the total mass of matched and unmatched inventors  $\nu M$ , with a search efficiency parameter  $\alpha$  and an elasticity with respect to search effort  $\theta$ . I denote  $dF^N(\hat{z})$  as the mass of unmatched divisions with relative productivity  $\hat{z}$ , and  $h$  as their search effort; and I denote  $dF^V(\hat{z}, q)$  as the mass of matched divisions with relative productivity  $\hat{z}$  and match quality  $q$ , and  $h_q$  as their search effort. Note that, because innovation is unrelated to  $\hat{z}$ , so are the returns to search.

I denote  $m^*$  and  $m^*(q)$  as the optimal search effort of unmatched divisions and matched divisions with match quality  $q$ , respectively. I am now ready to write out the aggregate meeting rate  $H$

---

<sup>16</sup>Assuming diminishing returns is analogous to introducing convex vacancy posting costs in a standard search model. As [Kaas and Kircher \(2015\)](#) show, this is empirically relevant, and it simplifies the model computationally. Also, without such an assumption, matched divisions would never search. With a linear search technology, unmatched divisions would be exactly indifferent between searching and not searching, and because matched divisions can set  $r$  to an arbitrarily small value they are strictly better off than if they were unmatched.

$$H = \left( \underbrace{\alpha \int_{\hat{z}} m^* dF^N(\hat{z}) + \int_{\hat{z}, q} m^*(q) dF^V(\hat{z}, q)}_{\text{Search Effort } S} \right)^\theta (\nu M)^{1-\theta}$$

I next introduce some helpful notation. First, it will be useful to write out the “contact rates”, or the number of meetings per unit of search effort and the number of meetings per inventor, which I respectively denote  $C_s$  and  $C_i$ :

$$C_s = \frac{H}{S}, \quad C_i = \frac{H}{\nu M}$$

Upon meeting, the division and inventor draw a match quality  $q$  from a uniform distribution with support  $[q_l, q_h]$ . They then decide whether to break any current match that they have and match with each other. If both agree to the match, the match retains its quality  $q$  throughout its duration.

Before writing out the firms’ problem of choosing its search effort, I introduce the notation  $\Omega$  for the expected value of meeting a new inventor. It is as follows for an unmatched division:

$$\begin{aligned} \Omega = & \underbrace{\left( 1 - \int_{\hat{z}', q'} \frac{dF^V(\hat{z}', q')}{\nu M} \right) \left( \int_q (V(\hat{z}, q) - N(\hat{z})) dF^q(q) \right)}_{\text{Meeting Unmatched Inventor}} \\ & + \underbrace{\int_{\hat{z}', q'} \int_{q > q'} (V(\hat{z}, q) - N(\hat{z})) dF^q(q) \frac{dF^V(\hat{z}', q')}{\nu M}}_{\text{Meeting Matched Inventor}} \end{aligned}$$

For a matched division with states  $\hat{z}, \tilde{q}$ , this instead becomes:

$$\begin{aligned} \Omega(q) = & \underbrace{\left( 1 - \int_{\hat{z}', q'} \frac{dF^V(\hat{z}', q')}{\nu M} \right) \left( \int_{q > \tilde{q}} (V(\hat{z}, q) - V(\hat{z}, \tilde{q})) dF^q(q) \right)}_{\text{Meeting Unmatched Inventor}} \\ & + \underbrace{\int_{\hat{z}', q'} \int_{q > \max[q', \tilde{q}]} (V(\hat{z}, q) - V(\hat{z}, \tilde{q})) dF^q(q) \frac{dF^V(\hat{z}', q')}{\nu M}}_{\text{Meeting Matched Inventor}} \end{aligned}$$

Divisions balance the costs of hiring labor for search against the benefits of making new matches. In equilibrium, labor hired for headhunting must be paid the same market-clearing wage  $w$  as

workers used in final production, and the returns to search (the benefit of meetings  $\Omega$  times meetings per unit of search  $H_s$ ) must equal their cost. For unmatched and matched divisions, respectively, the problems of choosing effort are:

$$\max_h m(h)\Omega C_s - wh$$

$$\max_h m(h)\Omega(q)C_s - wh$$

Matches break at an exogenous rate  $\delta_m$ .

## 1.6 Hamilton-Jacobi-Bellman Equations

In what follows I consider a transformed version of the economy on a balanced growth path: wages  $w$  and average productivity  $Z$  grow at a constant rate  $g$ , the distribution of divisions across relative productivities  $\hat{z}$  and match qualities  $q$  is constant, and policies are constant.

The value of an unmatched division has five parts: flow profit from its product, the possibility of creative destruction, the capital gains from drifting down in “relative productivity”, the gains from meeting inventors  $\Omega$ , and the costs of generating those meetings.

$$\rho N(\hat{z}) = \max_h \underbrace{\pi(\hat{z})}_{\text{Flow Profit}} - \underbrace{\Delta N(\hat{z})}_{\text{Creative Destruction}} - \underbrace{g(\sigma - 1) \frac{\partial N(\hat{z})}{\partial \hat{z}}}_{\text{Drift Down}} + \underbrace{m(h)\Omega C_s - wh}_{\text{Search}}$$

The value of a matched division has six parts: flow profits, the risk that the match breaks, the possibility of creative destruction, drifting down, innovation, and the risk that their inventor will meet another division.

$$\begin{aligned} \rho V(\hat{z}, q) = \max_{r, h} & \underbrace{\pi(\hat{z})}_{\text{Flow Profit}} - \underbrace{(\delta_m + \delta_i)(N(\hat{z}) - V(\hat{z}, q))}_{\text{Match Break}} - \underbrace{\Delta V(\hat{z}, q)}_{\text{Creative Destruction}} - \underbrace{g(\sigma - 1) \frac{\partial V(\hat{z}, q)}{\partial \hat{z}}}_{\text{Drift Down}} \\ & + \underbrace{m(h)\Omega(q)C_s - wh}_{\text{Search}} + \underbrace{(1 - \chi)qr^\gamma CD - wr}_{\text{Innovation}} \end{aligned}$$

$$+ \underbrace{\int_{q' > q} C_i (N(\hat{z}) - V(\hat{z}, q)) \frac{\int_{\hat{z}} h^\eta dF^N(\hat{z}) + \int_{\hat{z}} \int_q^{q'} h_{\hat{q}}^\eta dF^V(\hat{z}, \tilde{q})}{S} dF^q(q')}_{\text{Inventor Finds New Division}}$$

## 1.7 Equilibrium

A balanced growth path consists of detrended values  $V, N$ , policies  $e, h, r$ , wages  $w$ , distributions  $F, F^V, F^N$ , and a growth rate  $g$  such that divisions maximize profits, workers maximize their utilities, and policies are consistent with the distributions and growth rate being constant over time.

## 1.8 Discussion of Assumptions

The stark, simplifying assumptions that I have imposed on inventor compensation and matching (no earnings when not matched, a fixed bargaining weight  $\chi$ , and equal meeting rates for employed and unemployed) make the model far more tractable. Crucially, they allow me to solve the model without solving for the inventors' value functions. To show this, I begin with the following helpful proposition:

**Proposition 1.** *Assume  $\gamma$  (the curvature of the innovation cost function) and  $\chi$  (the inventor's bargaining weight) are both strictly between 0 and 1. Further, assume that inventors and divisions only observe their own current and potential match quality  $q$ , and do not observe whether the other division or inventor is already matched. If an inventor and division meet and draw potential match quality  $q$ , and if  $\gamma$  strictly less than 1:*

1. *If both are unmatched, the meeting converts to a match.*
2. *If only one party is matched, the meeting converts to a match if and only if the new match quality  $q$  is higher than the quality of the party's current match.*
3. *If both are matched, the meeting converts to a match if and only if the new match quality  $q$  is greater than the quality of both party's current matches.*

*Proof.* First, recall that inventors have no outside option, and that the search technology is equally good for matched or unmatched inventors and divisions. A matched division has the option to set  $r = 0$  and search as intensely as if it was unmatched, and it has the option to accept and reject

the same potential matches as if it was unmatched. Hence, being matched is at least weakly better than being unmatched for a division. We can go further and show that this dominance must be strict: if  $0 < \gamma < 1$ , then as  $r \rightarrow 0$  the division's profits from innovation  $(1 - \chi)qr^\gamma CD - wr$  become positive. Hence, a matched division can do strictly better than an unmatched one by keeping the same search-and-match policy but engaging in a small amount of research.

Because a matched division will always engage in a positive amount of research, it will always pay out a positive amount in compensation to the inventor in expectation. A matched inventor can do strictly better than an unmatched one by accepting and rejecting the same future offers but taking the division's compensation in the meantime. Therefore, both divisions and inventors strictly prefer being matched to being unmatched.

Because both the argmax and maximum of  $(1 - \chi)qr^\gamma CD - wr$  are increasing in  $q$ , divisions and inventors both strictly prefer high  $q$  to low  $q$ . Hence, they will leave a lower- $q$  match if and only if a higher- $q$  match is available.

□

Recall that the *only* decision inventors need to make is whether to accept a match or not, and the *only* source of welfare in this model is consumption (there are no taste shocks or moving costs.) Hence, characterizing the acceptance sets of both the firms and workers using Proposition 1 allows me to solve and estimate the model without solving for worker value functions.

## 2 Data

In order to discipline my model, I target a series of novel moments about inventor-firm dynamics. My main source of data is PatentsView, the US Patent Office's digitized repository of all US patent applications since 1976. A major value-added of PatentsView relative to the raw patent data is their effort to disambiguate the names of inventors and the individuals or organizations to whom the patents are assigned: the raw data only includes alphanumeric names rather than a consistent identifier. I take PatentsView's disambiguation as a given and note that my results will be affected to some degree any time that two inventors are incorrectly misclassified as the same person or one inventor is wrongly classified as two.



The key patent-level variables that I use are the inventor, the assignee (the firms and individuals who were assigned ownership rights to the patent), and the filing date of the patent (I use filing rather than grant dates as the former are presumably closer to when the patent was created.)

I subset the data to only include US-based inventors, and I only focus on patents with at least one corporate assignee. I therefore do not need to be concerned with whether non-US patents or “backyard” patents are comparable to US corporate patents. Excluding “backyard” innovation also means that my dataset is more comparable to my model, which does not have such innovations. My sample ultimately includes 3.3 million unique patents, 1.4 million unique inventors, and 200,000 unique assignees.

## 2.1 Key Moments

In this section I introduce some of the moments that I use to discipline my model. In particular, these moments help me discipline the parameters that govern the dynamics of the market for inventors: the efficiency of matching  $\alpha$ , the share of the population that can be an inventor  $\nu$ , the rate of matches breaking  $\delta_m$ , the rate at which inventors lose their creativity and become non-inventors  $\delta_i$ , and the gap between the best and worst matches  $q_h/q_l$ .

Inventors in my sample have, on average, 5.66 patents across 1.6 assignees, and their first and last patents are separated by an average of 5.7 years.<sup>17</sup> This is consistent with a move once every 3.5 inventor-years or so, indicating about as much mobility among inventors as among the general population.<sup>18</sup>

As I do not have access to the actual dates of employment for each inventor, I cannot directly measure job-to-job or employment-unemployment transitions. I must instead use the dates when their patents were filed. In some cases I find that inventors patent for firm A, then firm B, then A again in rapid succession. This suggests that an inventor’s final patents at their old firm may be filed after their first patents at their new firm. Therefore, I code an employee as having moved in year  $t$  if the plurality of their patents were assigned to a different firm than the last time when they had non-zero patents.<sup>19</sup>

<sup>17</sup>If an inventor patents once in 2014 and once in 2019, I record their “career length” as 5 years. If a patent has multiple corporate assignees, I only count one.

<sup>18</sup>[US Bureau of Labor Statistics \(2022\)](#) finds that Baby Boomers had 2.1 jobs from ages 45 to 54, and had many more job changes at younger ages.

<sup>19</sup>If an employee has one patent at Microsoft in 2012, then one patent at Apple in 2014, I code them as having moved

Table 1: Key Moments

Description	Value
Inventor career length (years)	5.7
Average patents per inventor	5.7
Average assignees per inventor	1.6
Average inventor-firm tenure (years)	1.8
Share of patents to new inventors	50
Share of patents to poached inventors	20

*\*Notes:* “Inventor career length” is the average number of years between an inventor’s first and last patent. “Average patents per inventor” is the average number of unique granted patents that each inventor has, with coauthored patents treated the same as solo-authored patents. “Average assignees per inventor” is the number of unique corporate assignees in the patents granted to the inventors. “Average inventor-firm tenure” is the average number of years between the inventor’s first and last patents at a given assignee. “Share of patents to new inventors” refers to the share of patents at time  $t$  awarded to inventors who did not patent in  $t - 1$ . “Share of patents to poached inventors” refers to the share of patents at time  $t$  awarded to inventors who had patented for a different firm in the last year before  $t$  that I saw them in the data. See Section 2.1 for details.

I compute two moments to help discipline the dynamics of the inventor labor market. First, as a proxy for transitions into unemployment, I calculate the average number of years between an inventor’s first and last patent with the same assignee.<sup>20</sup> I find that this average time gap is around 1.8 years. Second, as a proxy for job-to-job flows, I calculate the share of patents granted to inventors whose assignee (employer) at year  $t$  is different from their employer in the last year when they had a nonzero number of patents. I find that this share is about 20 percent. Analogously, 50 percent of patents in year  $t$  are awarded to inventors who had not patented in  $t - 1$ , including inventors who had never patented at all. I summarize these moments in Table 1.

## 2.2 Matched Pairs

I follow Prato (2022) and discipline the effects of mobility on innovation using a matched-pairs design, pairing inventors that move with similar inventors who do not move. My units of observation are inventor-years where the inventor files for a patent.<sup>21</sup> My treatment observations are years in which an inventor moved and had a patent, and my placebo observations are years in which inventors had a patent but did not move.<sup>22</sup> I calculate “moves” in the same way as in the

in 2014. If an employee has one patent for Apple in 2012, then one for Apple and two for Google in 2014, I still code them as having moved in 2014 even if their patent at Apple was filed later in 2014 than their two patents at Google.

<sup>20</sup>If an inventor only patents in a single year for that assignee, they count as a 0 towards this mean.

<sup>21</sup>In Prato (2022), the units of observation are inventors. The differences is that I do not restrict my sample to the first time that I observe an inventor move.

<sup>22</sup>I allow one inventor-year to act as a placebo for multiple treatments.

last section: I code an inventor as having “moved” in year  $t$  if the plurality of the patents they filed in that year were assigned to a different firm than in the last year when they had nonzero patents.

As in [Prato \(2022\)](#) I match treatment years to placebo years based on the cumulative number of patents that both inventors had at  $t - 1$ , the inventors’ first year with a patent, the current year, and the fact that both had non-zero patents in year  $t$ . The last condition is important as it means that inventors patent in both the treatment and placebo years—that is, I am comparing treatment years with at least 1 patent to placebo years with at least 1 patent. I am not comparing individuals who move (and who, by definition, continue to patent) with individuals who stay but may stop patenting. Aside from these matching conditions, my large sample size gives me the opportunity to use stricter matching conditions than [Prato \(2022\)](#). I therefore choose my placebos from the list of inventors that never change firms, and I also match treatment to placebo based on their most recent employer before  $t$  and on the exact number of patents that both filed in the year  $t - 1$ . Naturally, I also choose both treatment and placebo from the set of inventors who patent in more than one year. Ultimately, I am left with a sample of over 75,000 inventor-years with a move matched to a placebo.

Note that, because my placebos are chosen from the set of inventors who stay at the same firm and keep patenting after  $t - 1$ , the original firm must continue to patent after the treatment inventor leaves. This should exclude many instances when a “move” across firms was in fact a merger or acquisition.

All of my specifications report differences between the treatment  $T$  and placebo  $P$ , and all parameters are year indicators. Note that I do not include  $t - 1$  in these specifications, as I match the treatment and placebo on their exact number of patents in that period.

$$Y_j^T - Y_j^P = \sum_{\tau \in [-5, -2] \cup [0, 5]}^5 \beta_\tau \mathbb{1}(j = \tau) + \epsilon \quad (1)$$

I use four different outcomes for  $Y$ . In my first and primary specification,  $Y_j^i$  is the number of patents filed by  $i$  in relative year  $j$ . In my second specification,  $Y_j^i$  is instead the number of patents multiplied by their number of citations. In my third and fourth specifications, I ask whether my results are driven by inventors moving to systematically more-productive firms, rather than movement along a match quality “ladder” as in my model. To do so, I take the same treatment-placebo pairs and replace their own patent counts (both raw and citation-weighted) with the

average number of patents per inventor at the firms that they worked at.<sup>23</sup> I show the results of this analysis in Table 2: mobility is associated with a large increase in the treated inventor's own patents and patents-times-citations, but is not associated with an increase for the firm-level analogues of these variables. That is, inventors seem to become more productive when they move, but do not necessarily go to firms that systematically patent more.

Table 2: Coefficients

	Raw Patents	Citations	Firm Patents	Firm Citations
	(1)	(2)	(3)	(4)
$\beta_{-5}$	0.006 (0.005)	0.694 (0.376)	-0.081*** (0.002)	-1.503*** (0.174)
$\beta_{-4}$	-0.002 (0.005)	0.699* (0.308)	-0.068*** (0.002)	-0.876*** (0.142)
$\beta_{-3}$	-0.025*** (0.004)	0.166 (0.270)	-0.051*** (0.002)	-0.578*** (0.125)
$\beta_{-2}$	-0.035*** (0.003)	0.358 (0.224)	-0.026*** (0.001)	-0.400*** (0.070)
$\beta_0$	0.122*** (0.007)	8.898*** (0.456)	-0.158*** (0.003)	2.936*** (0.270)
$\beta_1$	0.243*** (0.008)	7.236*** (0.344)	-0.273*** (0.004)	-0.007 (0.232)
$\beta_2$	0.268*** (0.008)	7.002*** (0.386)	-0.255*** (0.004)	0.233 (0.220)
$\beta_3$	0.282*** (0.008)	6.728*** (0.307)	-0.243*** (0.004)	0.580** (0.202)
$\beta_4$	0.294*** (0.008)	6.939*** (0.376)	-0.242*** (0.004)	0.696*** (0.201)
$\beta_5$	0.303*** (0.009)	6.044*** (0.351)	-0.242*** (0.004)	0.635** (0.197)
Estimator	OLS	OLS	OLS	OLS
$N$	603,641	603,641	603,641	603,641
$R^2$	0.014	0.005	0.053	0.001

\* Notes: This table reports the results of Equation 1 for four different outcome variables  $Y$ : the inventors' own raw patent count, the inventors' count of patents-times-citations, and the average of the same two variables at the level of the firm. To calculate the firm-level variables, I replace the inventors' own patents or patents-times-citations with the average at the firm where they had most of their patents in that year. If they did not patent in a given year, I use the last firm I saw them patenting at.  $\beta_j$  is the coefficient on an indicator variable taking value 1 in year  $j$ , where  $j$  is defined relative to the treatment year.  $N$  refers to the number of observation-years, hence each treatment-placebo pair can appear up to 10 times (from  $j = -5$  to  $j = -2$ , and then from  $j = 0$  to  $j = 5$ ). Because I match the treatment and placebo on the exact number of patents in  $j = -1$ , I drop that year from this regression. I cluster standard errors at the level of the treatment-placebo pair.

<sup>23</sup>If an inventor patents at Apple in 2014 and Microsoft in 2016, I replace their patents with the average inventor-level patent count at Apple in 2014 and Microsoft in 2016, but keep them at 0 in 2015.

To help visualize and interpret the results more easily, I scale all coefficients and standard errors by the mean of  $Y_j^T$  for  $j \in [-5, -1]$  and plot them in Figure 1. A coefficient of 0.4 means that the difference in  $Y$  between the treatment and placebo at time  $j$  was 40 percent of the treatment's mean value of  $Y$  in the pre-treatment period.

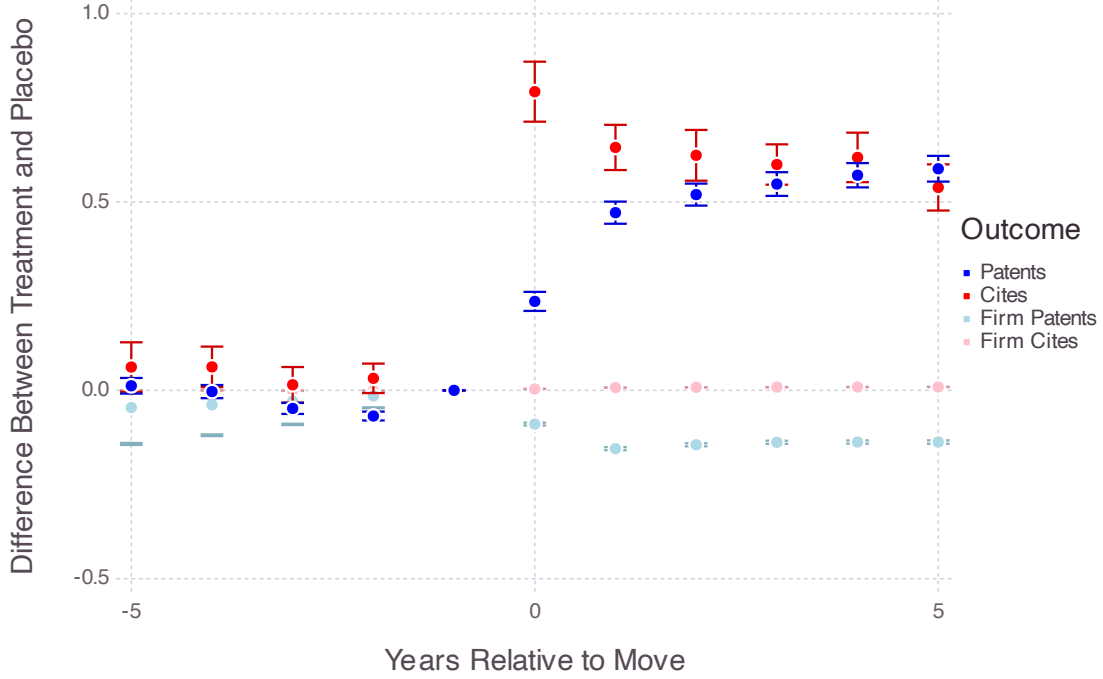


Figure 1: Difference in Outcomes, Treated vs Placebo

*Notes:* This figure illustrates the results of Equation 1 and Table 2. I again report the difference between the treatment and placebo inventors along four outcome variables  $Y$ : own patents, own patents-by-citations, and their firm-level counterparts. I scale the coefficients and standard errors by the mean value of  $Y$  for the treated inventors in the pre-treatment period. Hence, a value of 0.5 can be read as “the difference in  $Y$  between the treatment and placebo was 50% of the baseline value of  $Y$  for the treatment.” See the notes for Table 2 for additional details.

Treated inventors patent more than placebos after they move, a pattern that is robust to scaling up the raw count of patents by the number of forward citations that they receive. These patterns are not driven by the inventors moving to firms that systematically have more patents per inventor—if anything, the raw citation counts indicate that they move to firms that tend to have *fewer* patents per inventor. This result changes if I scale the raw firm-level patents by their forward citations, but those coefficients are essentially zeros.

I also find that my results are driven by an increase in patenting among treated inventors, rather than a fall in patenting among the placebos. To show this, I calculate the ratio of patents in the

pre-treatment period to patents in the post-treatment period for the placebo and treatment and demonstrate that both are well above 1:

$$\frac{\text{Mean Patents of Treatment, } t \geq 0}{\text{Mean Patents of Treatment, } t < 0} \sim 1.73$$

$$\frac{\text{Mean Patents of Placebo, } t \geq 0}{\text{Mean Patents of Placebo, } t < 0} \sim 1.18$$

To summarize these results in a single number (which I replicate with simulated data and target as a moment), I calculate the average change in the percentage difference in patents between the treatment and placebo as follows:

$$\frac{\text{Mean Patents of Treatment, } t \geq 0}{\text{Mean Patents of Placebo, } t \geq 0} - \frac{\text{Mean Patents of Treatment, } t < 0}{\text{Mean Patents of Placebo, } t < 0}$$

I find that this number is around 0.4—that is, the average treatment inventor has the same number of patents as the average placebo inventor in the five years before the move, but has 40 percent more patents in the six years during and after the move.

### 3 Quantitative Exercise

I now explain how I map the model to data, with the goal of quantifying the losses caused by search frictions in the market for inventors, asking how or whether policy can recover those losses, and testing how well the model can replicate key firm dynamics facts.

#### 3.1 Calibration

The model has fourteen parameters: the elasticity of substitution across varieties  $\sigma$ , the rate of discounting  $\rho$ , the innovation curvature parameter  $\gamma$ , the inventor bargaining weight  $\chi$ , the firm entry congestion curvature  $\psi$ , the elasticity of the meeting function  $\theta$ , the lower and upper bounds of the match quality distribution  $l$  and  $h$ , the share of the population  $\nu$  that can be an inventor, the exogenous match break rate  $\delta_m$ , the efficiency of the search technology  $\alpha$ , the step size  $\lambda$ , the efficiency  $\iota$  of the firm entry technology, and the measure of workers  $M$ .

### 3.1.1 Externally-Calibrated Parameters

I set the elasticity of substitution  $\sigma$ , discounting  $\rho$ , and curvature in the innovation production function  $\gamma$  to standard values of 0.05, 4, and 0.5, as in (for example) [Peters \(2020\)](#). I set the elasticity of the meeting technology  $\theta$  to a standard value of 0.5, as in [Fujita and Moscarini \(2017\)](#) or [Gavazza et al. \(2018\)](#). I follow [Peters and Zilibotti \(2021\)](#) in imposing a small degree of congestion ( $\psi = 0.1$ ) in the entry technology for firms. As in that paper, this is purely for computational convenience.<sup>24</sup> For the curvature of the match cost technology  $\eta$ , I follow [Kaas and Kircher \(2015\)](#) and set it to 0.5.

### 3.1.2 Internally-Calibrated Parameters

I match the remaining nine parameters by replicating key moments in both the data and the model. For a subset of moments, I calculate analogues using the steady state and do not simulate. For the remainder of these moments, I simulate a panel of 30,000 inventors over 40 years and calculate the same moments in the data as in the model. I estimate all parameters jointly but give a heuristic explanation of which moments map to which parameters in this section.

I scale the level of three parameters (the highest and lowest match qualities ( $q_h$  and  $q_l$ ) and the entrant productivity  $\iota$ ) inside the solution algorithm to ensure that GDP growth is 2 percent, and target the ratios of these three parameters with additional moments. I target  $q_h/q_l$ , the ratio of productivities of the most and least productive matches, so that the model can reproduce the 40 percent increase in patenting that I find with my matched-pairs design. I calculate the same moment in the simulated data as in the real data: I code an inventor as having moved in year  $t$  if the plurality of their patents in  $t$  were assigned to a different firm than in the last year when they had a patent. I use the level of the new-firm creation parameter  $\iota$  (or, more precisely, the ratio between  $\iota$  and  $q_l$ ) to target the share of employment in firms less than one year old, which was roughly 1.4 percent in 2018 according to [Hong and Werner \(2020\)](#). Because an entrant firm consists of a single division and operates a single product line at first, I approximate this moment without simulation by multiplying the mass of entrants by the average employment in a newly-created division.

I set  $\lambda$  to target an exit rate of “new” (less than 1 year old) firms of 14 percent, a value I take from [Peters and Walsh \(2022\)](#).<sup>25</sup>

---

<sup>24</sup>This ensures that the amount of labor demanded for entry will never have a corner solution.

<sup>25</sup>Growth in this model is equal to  $\lambda\Delta/(\sigma - 1)$ , and the exit rate of entrant firms (which start with a single product) is

To calibrate  $\nu$ , the exogenous share of workers who are capable of being an inventor, I target the share of workers who patent in a given year. In the data, I divide the number of individuals with a patent in each of the years from 1976 to 2022 by the size labor force in that year and take the average of that value. On average, fourteen in every ten thousand workers have a patent each year.

I calibrate the match break rate  $\delta_m$  to match the 1.8 years separating an inventor's first and last patent within the same firm.<sup>26</sup> To discipline the rate  $\delta_i$  at which inventors lose their creativity and become non-inventors, I target the average length of time between an inventor's first and last patent.

I use  $\alpha$ , the efficiency of the meeting technology, to target the 20 percent of patents in year  $t$  that were awarded to inventors who move to a new firm in  $t$ .<sup>27</sup> This is a rough proxy for the number of new matches in the economy and is therefore informative about the rate at which meetings happen.

To discipline the bargaining weight  $\chi$ , I use evidence collected in [Koh et al. \(2020\)](#) that labor comprises roughly 60 percent of the cost of investment in intellectual property. The closest analogue to these labor costs in the model are inventor salaries, as I interpret  $r$  as labor used to produce physical inputs for research. I therefore calculate the ratio of inventor pay to inventor pay plus the cost of  $r$  and target a value of 0.6.

I have normalized the measure of product lines to 1, but I cannot similarly normalize the measure of workers  $M$ . Because  $\nu M$  controls the number of inventors per product line, and because  $q$  is rescaled to ensure GDP growth is 2 percent,  $M$  helps the model rationalize the number of patents per inventor. That is, growth in this economy roughly depends on the number of inventors per product line times the average number of innovative steps per inventor.  $M$  scales the first of these two numbers so that the model can rationalize the second while still producing 2 percent GDP growth.

---

approximately  $\Delta$ . In order to reconcile 2 percent growth, an elasticity of substitution of 4, and a creative destruction rate of 14 percent,  $\lambda$  must be  $3/7$ . The fact that I am rescaling the values of the  $q$  grid and  $\iota$  to hit the growth rate means that the model will produce a creative destruction rate of almost exactly 14 percent if I fix  $\lambda$  to roughly  $3/7$ . Note that, because I solve the model on a discretized grid for  $\hat{z}$ , the "effective" value of  $\lambda$  will be slightly lower than  $3/7$ , as steps are bounded by the top of the grid, and so the creative destruction rate will be marginally higher than 14 percent. In practice, the inaccuracies that this introduces are tiny.

<sup>26</sup>As with the same moment in the data, an inventor who only patents in a single year counts as 0 towards this mean.

<sup>27</sup>I use the same approach as in the matched-pairs design: I count an employee as having moved if the plurality of their patents in year  $t$  were awarded to a different firm than the last year when they had a patent, and I do this in both in the real and simulated data.



### 3.1.3 Other Parameters

I add two additional “parameters” for the full quantitative implementation.

First, not every innovation is recorded as a patent, and my estimates will be biased if I map innovations directly onto patents. [Arundel and Kabla \(1998\)](#) find that 35 percent of innovations are recorded as patents in their sample of large European industrial firms, which I use as a benchmark. I therefore set  $\kappa = 0.35$  as the share of ideas that are patentable, and only record a random 35 percent of innovative steps as patents in my simulated data.

Second, I add an R&D subsidy, which I model as a fixed percentage subsidy  $s_r$  for all expenditures on  $r$  workers and compensation to inventors. The problem of a matched division choosing  $r$  becomes:

$$\max_r (1 - \chi(1 - s_r))qr^\gamma CD - wr(1 - s_r)$$

The government finances these subsidies in the model with taxes on the labor income of non-inventors. As non-inventors supply their labor inelastically, such taxes are non-distortionary, and I do not need to loop over taxes to balance the budget.

I treat  $s_r$ , the effective R&D subsidy, as a parameter to be estimated. [Tyson and Linden \(2012\)](#) report that in 2008, the R&D tax credit cost 11 billion dollars, and direct government support for corporate research cost another 26 billion. With US GDP at about 15 trillion that year — see [US Bureau of Economic Analysis \(2022\)](#) — I target a 0.0025 GDP share of corporate research subsidies.

## 3.2 Results

In Table 3 I present the results of the estimation. The model performs well on all targets, even though it has too few degrees of freedom to match all moments perfectly. Crucially, it does a good job of matching the increase in patenting that a “treated” inventor experiences upon a move.

In the model, 15.5 percent of GDP is paid out as investments in innovation (compensation to  $r$  or to inventors). [Atkeson and Burstein \(2019\)](#) report a wide range of reasonable estimates for the right GDP share of such investments, from the 3 percent of value added in research and development on the low end to the estimated 13 percent of non-farm output in the broader measure of [Corrado, Hulten, and Sichel \(2009\)](#), with their preferred value of 6 percent of value added splitting the

Table 3: Calibration

	Description	Value	Target	Data	Model
$q_l$	Lowest match quality	1.4	GDP growth	2%	2%
$q_h$	Highest match quality	1.73	Patenting change after move	40%	37%
			Average assignees per inventor	1.6	1.9
$\lambda$	Step size	0.43	Entrant exit	14%	14%
$\nu$	Inventor share	0.002	Share of workers with a patent	0.14%	0.14%
$\delta_m$	Exog. match break	1.1	Average inventor-firm "tenure" (years)	1.85	1.77
$\delta_i$	Inventor exit	0.12	Inventor "career" lengths (years)	5.7	5.5
			Average patents per inventor	5.7	5.6
$\alpha$	Search efficiency	9.4	Share of patents to "poached" inventors	0.2	0.22
			Share of patents to "new" inventors	0.5	0.52
$\chi$	Inventor bargaining weight	0.54	Inventor pay / research costs	0.6	0.63
$\iota$	Firm creation efficiency	0.03	Employment share of startups	1.4%	1.4%
$M$	Measure of workers	8.17	Average patents per year	1.0	1.0
$s_r$	Research subsidies	0.18	Subsidies/GDP	0.0025	0.0026
$\rho$	Discounting	0.05	Standard, i.e. <a href="#">Peters (2020)</a>		
$\sigma$	Elasticity of substitution	4	Standard, i.e. <a href="#">Peters (2020)</a>		
$\gamma$	Innovation cost curvature	0.5	Standard, i.e. <a href="#">Peters (2020)</a>		
$\alpha$	Meeting elasticity	0.5	Standard, i.e. <a href="#">Fujita and Moscarini (2017)</a>		
$\eta$	Search cost curvature	0.5	<a href="#">Kaas and Kircher (2015)</a>		
$\kappa$	Patentable share of ideas	0.35	<a href="#">Arundel and Kabla (1998)</a>		
$\psi$	Entry congestion	0.1	Set externally—see <a href="#">Peters and Zilibotti (2021)</a>		

\*Notes: This table reports the parameters for the calibrated model and compares model-generated moments with the data. All parameters affect all moments, but I list the parameters next to the moments which they most strongly affect. I take growth from the BLS, the rate at which new firms exit from [Peters and Walsh \(2022\)](#), the labor share of research costs from [Koh et al. \(2020\)](#), the entrant share of employment from [Hong and Werner \(2020\)](#), the share of GDP going to R&D subsidies from [Tyson and Linden \(2012\)](#) and [US Bureau of Economic Analysis \(2022\)](#). I calculate the rest of the internally-targeted moments from PatentsView.

difference. The model's inferred value of investments in innovation is towards the upper range of plausible estimates, but is not unreasonable. As there is significant uncertainty about the true share of output that acts as an investment in innovation, I leave this as an untargeted moment.

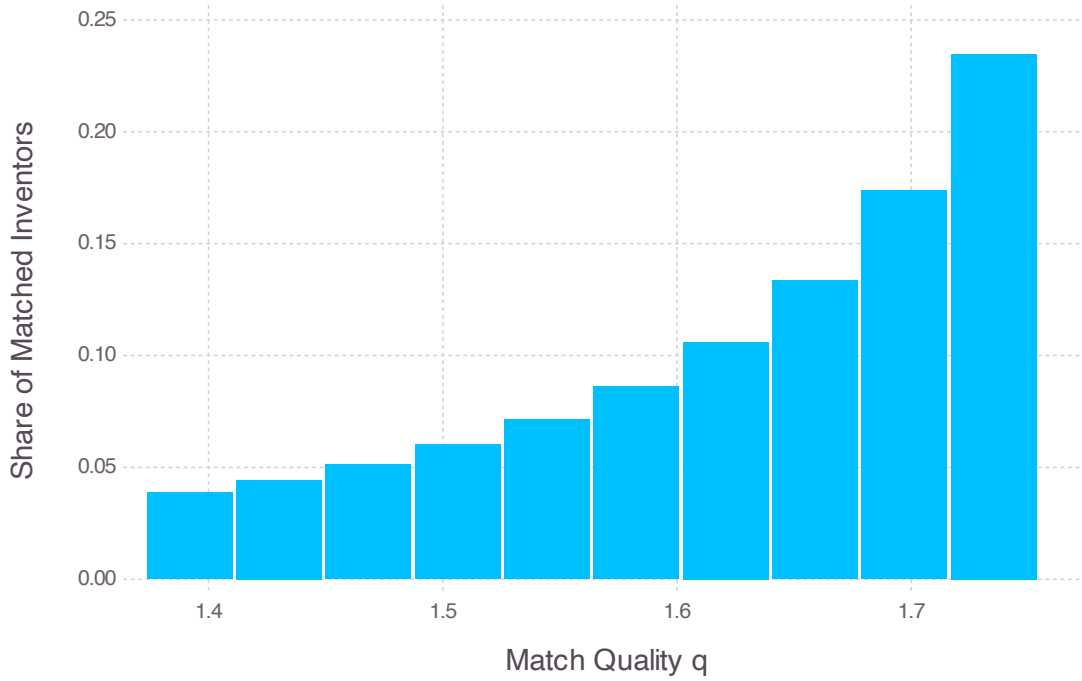


Figure 2: Distribution of Matched Inventors Across Match Qualities

*Notes:* This figure plots the share of matched inventors in each level of match quality. I discretize the model with 10 points on the match quality grid. Inventors and firms draw a match quality  $q$  when they meet, but because both inventors and firms look for better matches, they climb up the match quality ladder. Hence, almost a quarter of inventors have the best possible match quality, while less than 5 percent have the worst.

### 3.2.1 The Allocation of Inventors

64 percent of potential inventors are matched in equilibrium. This surprisingly low number is an artefact of my choice not to let them take other jobs when not working as an inventor.<sup>28</sup> Recall that the minimum value of  $q$  is around 1.5 and the maximum is around 1.7. The average match quality conditional on being matched is 1.6, indicating that inventors are well-matched in general. In Figure 2, I display the distribution of matched inventors across the match quality distribution. When an inventor and firm meet, they draw their match quality from a uniform distribution. Because there is both on-the-job and replacement search, inventors and firms climb the “match quality ladder” and replace worse matches with better ones. In Figure 3, I show how the rate of innovation  $\phi$  varies with match quality  $q$ . Because firms supply inventors with the research good  $r$ , and because the optimal value of  $r$  increases with match quality  $q$ , the rate of innovation  $\phi$  grows faster than  $q$ . Still,

<sup>28</sup>Without this assumption, it is possible that certain values of  $q$  would be unacceptably low for an unmatched inventor. This would make it challenging to separately identify  $\alpha$  and the lowest value that  $q$  could take.



Figure 3: Match Quality and the Rate of Innovation

*Notes:* This figure plots the match quality  $q$  and rate of innovation  $\phi$  for different match qualities. I rank match qualities from 1 to 10 on the horizontal axis, and plot their corresponding  $q$  and  $\phi$  on the vertical axis after rescaling by the lowest value of both of those quantities. The rate of innovation scales up with both  $q$  and the quantity of the research good  $r$  that firms provide inventors, which itself is increasing in  $q$ . Hence,  $\phi$  grows faster than  $q$ .

the best possible match only innovates about 55 percent more than the worst.

### 3.3 Implications for Firm Dynamics

This model offers new insights on the determinants of firm growth and innovation: most product lines never match with an inventor before being creatively destroyed, which means that the remaining product lines need to innovate rapidly to make up the difference. Most newborn firms never match more than a handful of their products to inventors, but the small handful that do then spin off many product lines that give them more chances to match with inventors down the line. Some newborn firms therefore die without having innovated at all, whereas a small handful expand rapidly.

To show how introducing search frictions can lead to dramatic dispersion in innovation across firms, I simulate a 30-year panel of newborn firms, born at random dates within that time window,

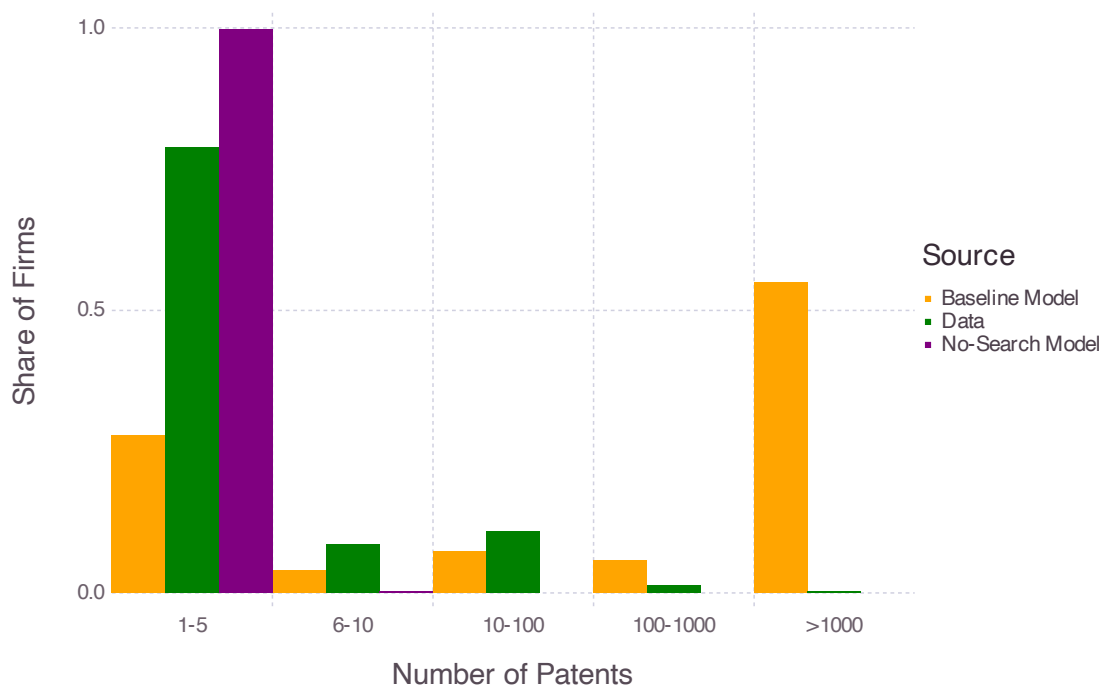


Figure 4: Distribution of Patents for “New” Firms

*Notes:* In this figure I plot the distribution of patent counts for entrant firms less than 30 years old in both model-generated and real data. I take the “Baseline Model” distribution from a simulation from the original calibrated model with a cap of 120,000 patents (the maximum in the data). I take the “No-Search Model” distribution from a simulation of a model where firms do not need to match with inventors before innovating. I take the “Data” distribution from the actual patent counts of firms that started patenting in 1990 or later. To ensure compatibility with the data, I draw the simulated firms’ birth years from a uniform distribution with support  $[0, 30]$ . Both simulations include 10,000 entrant firms, although not all firms successfully patent before their death. I bin firms into five groups based on their cumulative number of patents, and report the share of firms in each group on the vertical axis.

and tally up all the patents that these firms receive by the end of that window. To speed up the simulation, I do not allow firms to innovate after they receive 120,000 patents<sup>29</sup>, which is the maximum in the real data. In Figure 4, I compare the distribution of patents across entrant firms generated by the baseline model to the distribution generated by a “no-search” model where the intensity of innovation across product lines is the same, and to the distribution of patents awarded to firms that first patented in 1990 or later in the real data. In stark contrast to the “no-search” model, and indeed to the bulk of the Schumpeterian growth literature, the model with search frictions generates *more* dispersion in patenting than we see in the data for these young firms:

Recall that this model has no firm- or inventor- level permanent heterogeneity, only an idiosyn-

<sup>29</sup>Many firms actually hit the cap of 120,000 lifetime patents, which explains why the distribution appears to have a second mode at the far right end of the graph.

cratic match quality term. It does not need such features, or other modeling devices such as the “Luttmer Rocket”, to rationalize why some firms go from small startups to innovation powerhouses in a very short span of time. Note also that this is an untargeted moment—the model can generate at least as much dispersion in firm-level innovation as we see in the data without explicitly targeting it. Introducing search frictions for inventors to the Schumpeterian growth framework therefore helps make sense of some puzzling, hard-to-match firm dynamics facts without needing to “bake in” unexplained heterogeneity.

To help build more intuition for this result, and to understand what parameters are most relevant, note that in the standard framework of [Klette and Kortum \(2004\)](#) new divisions creatively destroy just under one rival division on average before being creatively destroyed themselves.<sup>30</sup> In the absence of any sort of firm heterogeneity, the exact number of “child” divisions follows a Poisson distribution, with a mean and variance of slightly less than 1. In my model, divisions need to match to an inventor before they innovate, and the odds of innovating before being creatively destroyed are quite low at just under 11 percent. Conditional on innovating at all, the number of innovations per division has a mean and variance of just over 7. Hence, introducing search frictions for inventors is a way of adding heterogeneity in growth rates, which especially drives dispersion for young firms with only a handful of inventors. A newborn firm with a single division that finds a good match quickly and survives for a long time can grow extremely rapidly. A fortunate match for such a firm therefore behaves similarly to an endogenous Luttmer Rocket.

The most important combination of parameters driving these results is  $M\nu$ , the number of workers per product line times the share of workers who can be inventors. At the current calibration, this number is extremely low at 0.02.<sup>31</sup> Very few divisions get to innovate, and those that do then innovate rapidly.

---

<sup>30</sup>Without entrant firms, which are a relatively small source of growth, new divisions would need to creatively destroy exactly one rival on average to keep the mass of product lines fixed at 1.

<sup>31</sup>Note that inventors cycle in and out of the economy at a high rate, and matches tend to break very quickly. This helps explain why 11 percent of divisions get to innovate before being destroyed even though only 2 percent can be matched at any one time.

## 4 Counterfactuals

In the next sections I study the costs of the misallocation of inventors, and ask whether policy can recover some of these costs.

### 4.1 The Cost of Search Frictions

In order to establish a benchmark or upper bound for the cost of the misallocation of inventors, I begin by increasing the efficiency of the search technology  $\alpha$  until the average match quality  $q$  of a matched inventor gets close to the highest possible value and until almost all inventors are matched. This should not be interpreted as a policy counterfactual but as an upper bound for what policy can possibly do. I find that growth increases from 2.0 to 2.6 percent as search frictions disappear, an enormous gain with significant consequences for welfare. In Figure 5, I plot the growth rate, share of matched inventors, average match quality conditional on being matched, share of labor allocated to producing the research good  $r$ , and consumption equivalent welfare (all normalized to 1 in the baseline economy) as  $\alpha$  increases and search frictions disappear. The dashed horizontal lines indicate the upper bounds of match quality and the matched share of inventors.

The share of labor allocated to producing the research good  $r$  does not scale up as quickly as the share of inventors who are matched. This is because a greater rate of innovation also means a greater rate of creative destruction, which lowers the returns to innovative effort and lowers the research intensity for a given inventor-firm pair.

Even though reducing search frictions in the market for inventors would significantly enhance growth, it does not follow that the economy under-invests in inventor search  $h$  at the current value of  $\alpha$ , or that there is a useful role for policy. The economy faces a simple dynamic tradeoff: On the one hand, a unit of labor that is allocated to the meeting technology increases the rate at which firms match with inventors, which will increase the average match quality  $q$  and therefore the growth rate of technology and consumption. Because of the quality-ladder structure of innovation, firms do not fully internalize the benefits of making a good match and innovating more quickly: some of the benefit goes to the rival firms that will creatively destroy their product lines. On the other hand, the labor used in the meeting technology cannot be allocated to producing the final good, so there is less consumption holding fixed the level of technology. Because there is congestion

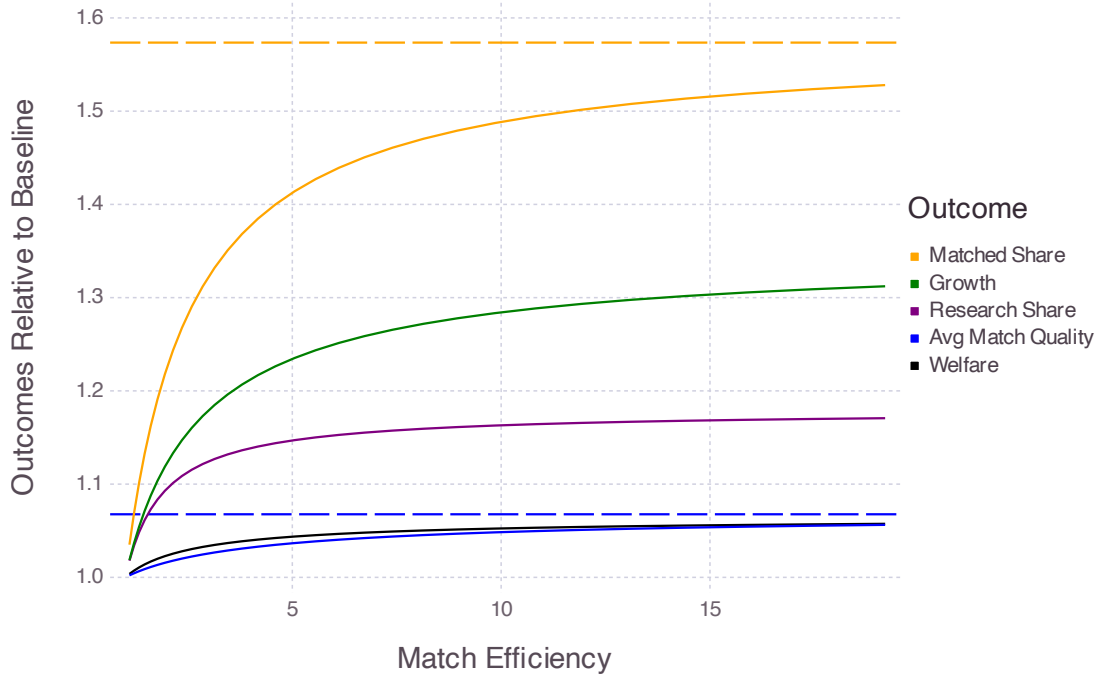


Figure 5: The Impact of Improving Match Efficiency

*Notes:* This figure plots the growth rate, share of matched inventors, average match quality conditional on being matched, share of labor allocated to producing the research good  $r$ , and consumption equivalent welfare as  $\alpha$ , the efficiency of the meeting technology, increases. To plot variables on the same scale and to make interpretation easier, I normalize all quantities so that they are equal to 1.0 at the baseline calibration. Hence, a value of 1.01 corresponds to a 1 percent increase. The horizontal axis is the efficiency of the meeting technology, also normalized so that the baseline is equal to 1. The dashed horizontal lines indicate the upper bound of match quality and the matched share, and they have the same colors as the variables they map onto.

in the market for inventors, the resources expended in search may even be too high. In order to understand whether knowledge or congestion spillovers dominate, I introduce a “wedge”  $\tau$  into the firm’s decision problem for search effort and re-solve for the new distorted equilibrium. A positive wedge  $\tau > 0$  induces firms to spend more than is privately optimal, and a negative wedge  $\tau < 0$  induces them to spend less. Note that these wedges leave the efficiency of search unchanged. The firms now maximize the following distorted problem:

$$\max_h m(h)\Omega H_s - wh(1 - \tau).$$

In Figure 6, I plot the rate of growth, consumption-equivalent change in welfare, share of matched inventors, and match quality conditional on being matched as I increase and decrease



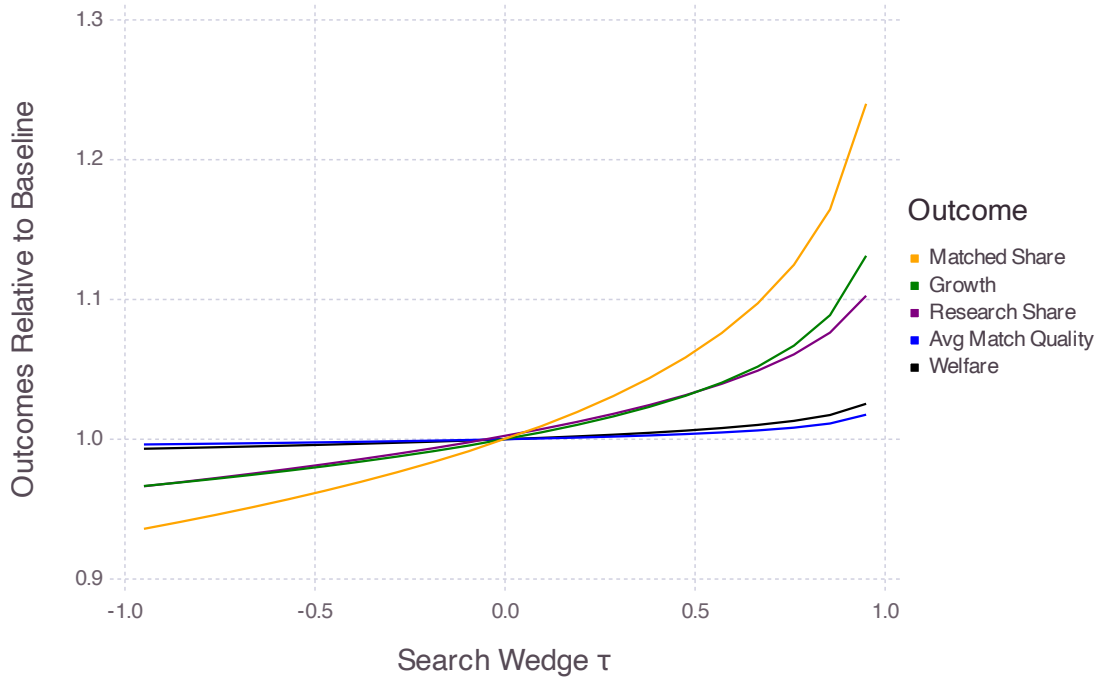


Figure 6: The Impact of Changing Search Effort

*Notes:* This figure plots the growth rate, share of matched inventors, average match quality conditional on being matched, share of labor allocated to producing the research good  $r$ , and consumption equivalent welfare under different values of the distortion  $\tau$ . Firms decide on their search effort “as if” the price of search was  $(1 - \tau)w$  rather than  $w$ , meaning that positive values of  $\tau$  cause firms to increase their search effort. The efficiency of search  $\alpha$  is the same—this exercise asks whether, at the current level of search efficiency, firms invest too much or too little in search. To plot variables on the same scale and to make interpretation easier, I normalize all quantities so that they are equal to 1.0 at the baseline calibration.

the share of labor allocated to search for inventors. I find that increasing search effort would be welfare-improving, because the positive knowledge spillovers from more innovation through better-matched inventors dominate the negative spillovers from congestion.<sup>32</sup> I find that welfare is monotone-increasing for values of  $\tau$  quite close to 1. This is in part because my calibrated search costs are extremely low and the share of the workforce allocated to search is tiny, so even increasing the number of workers devoted to search many times over only reallocates a small share of the overall workforce.

<sup>32</sup>As an aside, I find that actually implementing these wedges as subsidies, even for values of  $\tau$  close to 1.0, costs the government far less than traditional R&D subsidies for a given increase in growth. This connects my paper to a long tradition in the search literature of considering whether there is “too much” or “too little” vacancy posting, dating back to the seminal contribution of Hosios (1990). I leave it to future work to figure out whether or how such subsidies could be implemented.

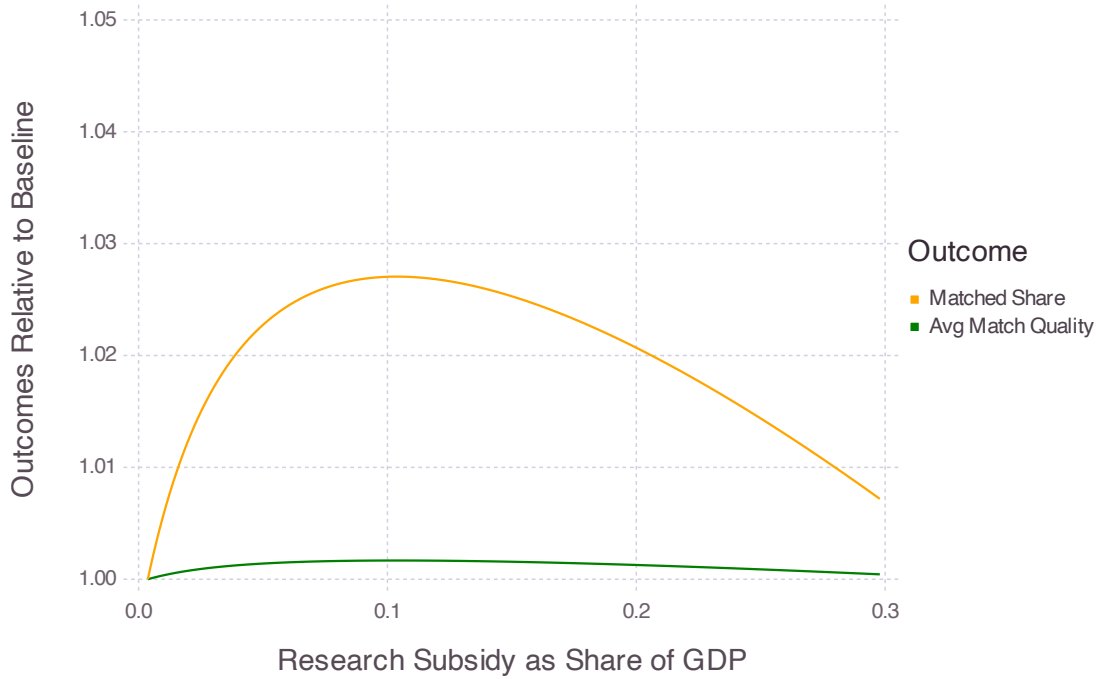


Figure 7: The Impact of R&D Subsidies  $s_r$  on Match Quality and the Share of Matched Inventors

*Notes:* This figure plots the share of inventors who are matched and their average match quality conditional on being matched as R&D subsidies increase. To plot both variables on the same scale and to make interpretation easier, I normalize both quantities so that they are equal to 1.0 at the baseline calibration. Hence, a value of 1.01 corresponds to a 1 percent increase. The horizontal axis is the R&D subsidy as a share of GDP, so a value of 0.3 means that 30 percent of GDP is allocated to R&D subsidies.

## 4.2 Can Policy Recover These Losses?

This model has two margins of investment: firms can invest in their “match capital” by increasing  $h$ , and they can invest in their “idea capital” by increasing  $r$ . I summarize the various subsidies to the latter form of investment as  $s_r$ , a stand-in for the R&D tax credits and other research subsidies. One might hope that R&D subsidies could bring the economy fairly close to the optimum by themselves: subsidizing  $r$  and inventors’ salaries increases the returns to search effort, and can in principle indirectly raise the average match quality to the point where there would be minimal gains from any further improvements.

In practice, I find that the indirect effects of R&D on the allocative efficiency of inventors are largely canceled out by their indirect effect on creative destruction. This intuition points to a novel reason why the benefits of R&D subsidies may not scale up: the additional creative destruction

coming from more innovation by rival firms destroys “match capital” and lowers the innovative potential of the economy. In contrast to standard Schumpeterian growth models, such capital is costly to build up and is lost to the economy when a division becomes obsolete.

In Figure 7, I plot the share of matched inventors and the average match quality of inventors conditional on being matched at all for R&D subsidies going all the way to 30 percent of GDP. Growth increases due to the large increase in  $r$  but match quality and the share of matched inventors stay basically flat. The share of matched inventors peaks no more than 3 percent higher than the baseline, whereas the average match quality conditional on a match only rises by a fraction of a percent. The positive incentive effects win out for any reasonable level of R&D subsidies, but the effects are small. Hence, R&D subsidies do not recover the losses from search frictions in the market for inventors and leave growth on the table.

## 5 Conclusion

In this paper I study a key step in the process of innovation: the allocation of inventors to firms. I find that inventor mobility across firms is associated with a large increase in the rate of innovation, and that such moves are quite frequent. Intuitively, the fact that mobility is associated with such large gains suggests that the quality of the match between an inventor and a firm is an important determinant of the rate of innovation, and the fact that inventors move so often suggests that the market for inventors is frictional and that inventors do not necessarily find the right match for their talents right away. I formalize these intuitions with a novel Schumpeterian growth model that features frictional search for inventors and a crucial role for the quality of the match between the inventor and the firm. I discipline this model with new moments about inventor-firm dynamics, and I use the model to shed light on the role that the market for inventors plays in firm-level and economy-wide growth. I show that search frictions for inventors can help rationalize why some firms grow dramatically quickly whereas others stagnate. I then show that search frictions in the market for inventors are a significant drain on growth, and that conventional innovation policies do not recover this lost growth. This paper opens up the black box of research spending and takes the market for inventors seriously, and shows that doing so helps open up new avenues for research into growth, firm dynamics, and innovation policy.

## References

- Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr (2018). Innovation, reallocation, and growth. *American Economic Review* 108(11), 3450–91.
- Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi (2018). Dancing with the stars: Innovation through interactions. Technical report, National Bureau of Economic Research.
- Arkolakis, C., M. Peters, and S. Lee (2020). European immigrants and the united states’ rise to the technological frontier.
- Arundel, A. and I. Kabla (1998). What percentage of innovations are patented? empirical estimates for european firms. *Research Policy* 27(2), 127–141.
- Atkeson, A. and A. Burstein (2019). Aggregate implications of innovation policy. *Journal of Political Economy* 127(6), 2625–2683.
- Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019). Who becomes an inventor in america? the importance of exposure to innovation. *The Quarterly Journal of Economics* 134(2), 647–713.
- Celik, M. (2022). Does the cream always rise to the top? the misallocation of talent in innovation. *Journal of Monetary Economics* (forthcoming).
- Celik, M. A. and X. Tian (2022). Agency frictions, managerial compensation, and disruptive innovations.
- Corrado, C., C. Hulten, and D. Sichel (2009). Intangible capital and us economic growth. *Review of income and wealth* 55(3), 661–685.
- Engbom, N. (2020). Misallocative growth. Manuscript.
- Fujita, S. and G. Moscarini (2017). Recall and unemployment. *American Economic Review* 102(7), 3875–3916.
- Garcia-Macia, D., C.-T. Hsieh, and P. J. Klenow (2019). How destructive is innovation? *Econometrica* 87(5), 1507–1541.

- Gavazza, A., S. Mongey, and G. L. Violante (2018). Aggregate recruiting intensity. *American Economic Review* 108(8), 2088–2127.
- Hong, S. and D. Werner (2020). Trends in Startups? Share of Jobs in the U.S. and Eighth District. *The Regional Economist* 28(1).
- Hosios, A. J. (1990, 04). On The Efficiency of Matching and Related Models of Search and Unemployment. *The Review of Economic Studies* 57(2), 279–298.
- Jones, C. I. and J. Kim (2018). A schumpeterian model of top income inequality. *Journal of Political Economy* 126(5), 1785–1826.
- Kaas, L. and P. Kircher (2015). Efficient firm dynamics in a frictional labor market. *American Economic Review* 105(10), 3030–60.
- Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Econometrica* 112(5), 986–1018.
- Koh, D., R. Santaaulàlia-Llopis, and Y. Zheng (2020). Labor share decline and intellectual property products capital. *Econometrica* 88(6), 2609–2628.
- Lentz, R. and D. T. Mortensen (2012). Labor market friction, firm heterogeneity, and aggregate employment and productivity. *Unpublished Manuscript*.
- Luttmer, E. G. (2011). On the mechanics of firm growth. *The Review of Economic Studies* 78(3), 1042–1068.
- Pearce, J. (2019). Firm-Inventor Links and the Composition of Innovation.
- Peters, M. (2020). Heterogeneous markups, growth, and endogenous misallocation. *Econometrica* 88(5), 2037–2073.
- Peters, M. and C. Walsh (2022). Population growth and firm-product dynamics. *Manuscript*.
- Peters, M. and F. Zilibotti (2021). Creative destruction, distance to frontier, and economic development. *The Economics of Creative Destruction - A Festschrift in Honor of Philippe Aghion and Peter Howitt*.

- Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica* 70(6), 2295–2350.
- Prato, M. (2022). The global race for talent: Brain drain, knowledge transfer, and economic growth.
- Shimer, R. (2006). On-the-job search and strategic bargaining. *European Economic Review* 50(4), 811–830.
- Tyson, L. and G. Linden (2012). The corporate r&d tax credit and us innovation and competitiveness: Gauging the economic and fiscal effectiveness of the credit. *Center for American Progress*, 21–2.
- US Bureau of Economic Analysis (2022). Gross domestic product. retrieved from FRED, Federal Reserve Bank of St. Louis; <https://fred.stlouisfed.org/series/GDP>.
- US Bureau of Labor Statistics (2022). Number of jobs, labor market experience, marital status, and health: Results from a national longitudinal study. retrieved from BLS; <https://www.bls.gov/news.release/pdf/nlsoy.pdf>.
- van der Wouden, F. and D. L. Rigby (2020). Inventor mobility and productivity: a long-run perspective. *Industry and Innovation* 28(6), 677–703.

# Appendices

## A Appendix: Solving the Model

I use a discretized version of the model, hence why I write sums rather than integrals in this section.

### A.1 Pseudocode

- Step 0: Initialize the model with a guess for distributions  $F, F^V, F^N$ ; general equilibrium objects  $\pi, g, \Delta, w$ ; values  $V, N$ ; and policies  $r, h, e$  as described below.
  - Re-scale the  $\hat{z}$  grid so  $Z = 1$ .
  - Guess that  $\min[\nu M \cdot 0.5, 0.5]$  inventors are matched, i.e. that either half of inventors or half of all divisions are matched (this ensures that, no matter what I put for  $M$  and  $\nu$ , I will not guess that more than a measure 1 of inventors or divisions are matched). Guess that matched divisions are uniformly distributed across the  $\hat{z}, q$  grid, and that unmatched divisions are uniformly distributed across the  $\hat{z}$  grid.
  - A reasonable first guess for  $V$  and  $N$ , conditional on the guess for the number of workers in production  $L^p$ , would be the present discounted value profits accruing to a division with productivity  $\hat{z}$  assuming no innovation, drift from growth, or creative destruction. I increase these values by 10% for  $V$  to take into account that matched division have the option to innovate, and I increase them by a further multiplicative factor equal to  $0.1 \cdot q/q_l$  for each value of  $q$  to take into account the greater value of a division with a good match.
  - Guess that 10 percent of non-inventor workers will work in producing the research good (this is  $r$ ).
  - Guess that 5 percent of non-inventor workers will work in entry (this is  $e$ ).
  - Guess that 0.1 percent of non-inventor workers will work in search (this is  $h$ ), and that search effort is uniform for all divisions.
- Step 1: Holding fixed the distributions, update policies, values, and GE objects.

- Step 1a: Holding fixed policies  $r, h, e$  and values  $V, N$ , solve for the GE terms  $\Delta, w, \pi, g$ .
- Step 1b: Re-scale every element of the  $q$  grid and  $\iota$  by a common term  $\kappa$  until  $g$  is 2%<sup>33</sup>, *unless* performing a counterfactual. Check if the growth rate had converged; i.e. check if it had already been 2% before rescaling.
- Step 1c: Holding fixed values and GE terms, solve for policies  $r, h, e$ .
  - \* If divisions demand  $R > 0.6M(1 - \nu)$  units of labor as an input into research, scale down their demands by a factor  $0.6M(1 - \nu) / R$  such that no more than 60% of the non-inventor labor force is used for research.
- Step 1d: Holding fixed policies and GE terms, update values  $V$  and  $N$  using an implicit step, and check if they have converged.
- Step 1e: Repeat 1a-1d until values and the growth rate converge, or until the maximum number of inner-loop iterations is reached.
- Step 2: Update the distributions using the policies recovered in Step 1.
  - If  $\sum_{\hat{z}, q} F^V(\hat{z}, q) > M\nu$ , i.e. there are more matched divisions than inventors, scale down the mass of matched divisions until only 90% of inventors are matched, and scale up the mass of unmatched divisions until the overall mass is again 1.0.
- Step 3: Repeat 1 and 2 until values, distributions, and the growth rate converge, or until the maximum number of outer-loop iterations is reached.

## A.2 Solving for GE objects

To calculate the rate of creative destruction  $\Delta$ , I sum up innovative activity among matched divisions and entrants.

$$\Delta = e\iota / E^\psi + \int_{\hat{z}, q} \phi(q) dF^V(\hat{z}, q)$$

This allows me to get the growth rate  $g$ , which is  $\Delta\lambda / (\sigma - 1)$ .

---

<sup>33</sup>I follow [Engbom \(2020\)](#) in holding fixed the growth rate and scaling the parameters to ensure that this growth rate is attained. I find, as he did, that the solution algorithm is more stable if I let the parameters vary but keep the growth rate fixed than vice versa.



Note that both  $\Delta$  and  $g$  can be scaled if we multiply  $\iota$  and every value of  $q$  by a multiplicative  $\kappa = 0.02/g$ . Doing so would rescale innovation such that the growth rate would be 2 percent if policies were held fixed.

Wages and flow profits are as usual in CES models of monopolistic competition:

$$w = Y/(\mu l)$$

$$\pi = (\mu - 1)\hat{z}wL$$

### A.3 Updating Policies

Solving for entry  $e$  is straightforward conditional on knowing the value of creative destruction  $CD = \sum N(\lambda \hat{z})F(\hat{z})$ . Recall the three key equations for this problem:

$$\phi_s = \iota/E^\psi$$

$$\phi_s CD = w$$

$$\phi_s e = E$$

Combining the first and second equation immediately yield:

$$E = \left( \frac{\iota CD}{w} \right)^{\frac{1}{\psi}}$$

Combining this with the first and third yields:

$$e = \frac{1}{\iota} \left( \frac{\iota CD}{w} \right)^{\frac{1-\psi}{\psi}}$$

Now I turn to the problem of the matched division choosing  $r$ . Conditional on  $CD$ ,  $w$ , and  $\tau_r$ , the division's effort level  $r$  can be solved in closed-form:

$$\max_r (1 - (1 - \tau_r)\chi)qr^\gamma CD - (1 - \tau_r)wr$$

$$r = \left( \frac{(\gamma(1 - (1 - \tau_r)\chi)qCD)}{(1 - \tau_r)w} \right)^{\frac{1}{1-\gamma}}$$

Finally, I calculate the number of workers hired as headhunters. Conditional on the value of meeting  $\Omega$  and on the contact rate per “unit of matching”  $C_s$ , search effort can be solved as follows for unmatched and matched firms, respectively:

$$\eta h^{\eta-1} \Omega C_s = w(1 - s_h)$$

$$h = \left( \frac{w(1 - s_h)}{\eta \Omega C_s} \right)^{\frac{1}{\eta-1}}$$

$$\eta h(q)^{\eta-1} \Omega(q) C_s = w(1 - s_h)$$

$$h(q) = \left( \frac{w(1 - s_h)}{\eta \Omega(q) C_s} \right)^{\frac{1}{\eta-1}}$$

With  $r, h, e$  solved, labor used in production  $l$  is  $M(1 - \nu)$  minus labor used for all three other functions.

#### A.4 Updating Values

I use a discretized version of the HJB equations, with  $\hat{z}$  and  $q$  on grids. I approximate the derivative of  $V$  and  $N$  with respect to  $\hat{z}$  using a standard upwinding approach:

$$\frac{\partial V(\hat{z}_n)}{\partial \hat{z}} \sim \frac{V(\hat{z}_n) - V(\hat{z}_{n-1})}{\hat{z}_n - \hat{z}_{n-1}}$$

I stack  $V$  and  $N$  into a single vector  $S$  and construct an intensity matrix across states  $A$  that summarizes all the flows across these states. This matrix takes policies as given and incorporates the costs of taxation and wages incurred when a division or inventor moves across states. I denote  $S$  as the stacked vector of values and  $f$  as the flow utility coming from profits, research costs, and effort.

$$S_{t+\epsilon} = S_t + \epsilon(f + AS_t - (\rho - g)S_{t+\epsilon})$$

This can be rewritten as:

$$\underbrace{\left[ \left( \frac{1}{\epsilon} + \rho - g \right) I - A \right]}_T S_{t+\epsilon} = \frac{1}{\epsilon} S_t + f$$

Finally we get the updated vector of values:

$$S_{t+\epsilon} = T^{-1} \left( \frac{1}{\epsilon} S_t + f \right)$$

I set a large value for  $\epsilon$ , check for convergence, and *then* update the value of  $S$  by taking a convex combination of  $S_{t+\epsilon}$  and  $S_t$ .

## A.5 Updating Discretized Kolmogorov Forward Equations

It is sufficient to characterize the distribution of divisions across states  $(\hat{z}, q)$ —knowing this distribution automatically pins down the distribution of workers. Denoting  $G$  as the stacked vector of distributions across states and  $K$  as the transition matrix across states, we get:

$$G_{t+\epsilon} = G_t - \epsilon K' G_t$$

It is possible to solve for the stationary distribution directly by iterating on this equation repeatedly or by looking for the solution to:

$$K' G = 0$$

I find the following approach works well: first, scale every element of  $K'$  by some factor  $s$  such that the sum of absolute values in each row is much less than 1.0 (recall that only the diagonal elements of  $K'$  are negative). Next, add  $I$  to  $K'$ . Then look for the solution to:

$$(I + sK')G = 0$$

$(I + sK')$  is a stochastic matrix where all elements are positive and the rows sum to 1, and I can find  $G$  using off-the-shelf routines for finding the stationary distribution of a stochastic process.

The matrix  $K$  used to update the distribution is different from the one used to update the values ( $A$ ). First, a creatively destroyed division does not derive value from the new division created to replace it, but the “destroying” division does. I do not keep track of the “destroying” distribution in  $A$ , but I do in  $K$ . Second,  $K$  does not incorporate taxes.

Notice one slightly unusual thing about  $K$ : the elements of the distribution actually show up directly in the matching probabilities. To solve for  $G$ , I hold  $K$  fixed and update till  $G$  converges, then I recalculate  $K$  and re-solve for  $G$ . I effectively iterate on solving for  $K$  and  $G$  until both converge.

## A.6 Extreme Values of $\hat{z}$

I solve the model on a discretized grid of  $\hat{z}$ , which does not have a natural upper or lower bound. In this section, I describe a few adjustments that I make to the model to deal with this discretization.

First, at the lowest point of the  $\hat{z}$  grid, I set the terms  $\partial V(\hat{z}, q) / \partial \hat{z}$  and  $\partial N(\hat{z}) / \partial \hat{z}$  to 0.

Second, I modify the value of creative destruction  $CD$ . If a division or unmatched inventor creatively destroys a rival division with  $\hat{z} > \hat{z}_{\max}$ , the highest value of  $\hat{z}$  on the grid, they spin off a new unmatched division with productivity  $\epsilon \rightarrow 0$  greater than the rival, not  $\lambda$  greater. Hence, the value of  $CD$  becomes:

$$CD = \sum N(\min[\hat{z}_{\max}, \lambda + \hat{z}])F(\hat{z})$$

Finally, when updating  $g$ , I set  $\Delta$  to 0 for the highest value of the grid.

## B Appendix: Deriving Growth and Drift Rates

In this section I offer heuristic explanations of how to calculate  $g$  and where the term  $g(\sigma - 1)\partial N(\hat{z})/\partial \hat{z}$  in the HJBs comes from. To do so, the following derivations about  $\hat{z}$  will be useful.

$$\hat{z} = \left(\frac{z}{Z}\right)^{\sigma-1}$$

$$Z = \left(\int_{\hat{z}} z^{\sigma-1} F(z)\right)^{\frac{1}{\sigma-1}}$$

$$Z = \left(\int_{\hat{z}} \hat{z} Z^{\sigma-1} F(z)\right)^{\frac{1}{\sigma-1}}$$

$$1 = \left(\int_{\hat{z}} \hat{z} F(z)\right)^{\frac{1}{\sigma-1}} = \int_{\hat{z}} \hat{z} F(z)$$

The growth rate  $g$  can be calculated by calculating the percent change in  $Z$  over an interval of time divided by the length of the interval, and then taking the limit of that expression as the length of the interval goes to zero.

$$g = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \frac{Z_{t+\epsilon} - Z_t}{Z_t}$$

Recall:

$$Z_t = \left(\int_z z^{\sigma-1} F_t(z)\right)^{\frac{1}{\sigma-1}} = Z \left(\int_{z_t} \frac{z^{\sigma-1}}{Z^{\sigma-1}} F_t(z)\right)^{\frac{1}{\sigma-1}} = \left(\int_z \hat{z} F_t(z)\right)^{\frac{1}{\sigma-1}}$$

At each point of the productivity grid, some share of the mass at that grid point jumps from  $\hat{z}$  to  $\lambda + \hat{z}$  at a rate of  $\Delta$ . Over a time step  $\epsilon$ , this means that:

$$Z_{t+\epsilon} = \left(\int_z (\hat{z} + \epsilon \Delta \lambda) F_t(z)\right)^{\frac{1}{\sigma-1}}$$

Dividing the numerator and denominator by  $Z_t$ , dropping time subscripts, and we get:

$$g = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \frac{\left(\int_z (\hat{z} + \epsilon \Delta \lambda) F_t(z)\right)^{\frac{1}{\sigma-1}} - 1}{1}$$

Simplifying a bit:

$$g = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \frac{(\int_z \hat{z} F_t(z) + \int_z \epsilon \Delta \lambda F_t(z))^{\frac{1}{\sigma-1}} - 1}{1} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \frac{(1 + \epsilon \Delta \lambda)^{\frac{1}{\sigma-1}} - 1}{1}$$

Taking the limit as  $\epsilon \rightarrow 0$  yields the expression in the paper:

$$g = \frac{\Delta \lambda}{\sigma - 1}$$

To understand where the term  $g(\sigma - 1)\partial N(\hat{z})/\partial \hat{z}$  in the value functions comes from, I work with a simplified discrete-time version of the model without innovation. This can be written as:

$$N(Z_t, \hat{z}_t) = \epsilon \pi(Z_t, \hat{z}_t) + e^{-r\epsilon} N(Z_{t+\epsilon}, \hat{z}_{t+\epsilon})$$

Because  $\pi$  scales multiplicatively in  $Z$  and  $\hat{z}$ , this can be rewritten with a constant term  $\tilde{N}$ :

$$\tilde{N} Z_t \hat{z}_t = \epsilon \pi Z_t \hat{z}_t + e^{-r\epsilon} \tilde{N} Z_{t+\epsilon} \hat{z}_{t+\epsilon}$$

Dividing both sides by  $Z_t$  and using the fact that  $Z_{t+\epsilon} = Z_t e^{\epsilon g}$ :

$$\tilde{N} \hat{z}_t = \epsilon \pi \hat{z}_t + e^{(g-r)\epsilon} \tilde{N} \hat{z}_{t+\epsilon}$$

Note that a standard Euler condition gives us that  $r = \rho + g$ , simplifying terms a bit. Subtracting  $\tilde{N} \hat{z}_t e^{-\rho\epsilon}$  from both sides and dividing by  $\epsilon$  we get:

$$\frac{\tilde{N} \hat{z}_t (1 - e^{-\rho\epsilon})}{\epsilon} = \pi \hat{z}_t + \frac{e^{-\rho\epsilon} \tilde{N} (\hat{z}_{t+\epsilon} - \hat{z}_t)}{\epsilon}$$

Taking the limit of both sides as  $\epsilon \rightarrow 0$  we get:

$$\rho \tilde{N} \hat{z}_t = \pi \hat{z}_t + \lim_{\epsilon \rightarrow 0} \frac{e^{-\rho\epsilon} \tilde{N} (\hat{z}_{t+\epsilon} - \hat{z}_t)}{\epsilon}$$

To simplify the right hand side, recall  $\hat{z}_t = (z_t / Z_t)^{\sigma-1}$ , meaning  $\hat{z}_{t+\epsilon} = (z_t / (Z_t e^{g\epsilon}))^{\sigma-1} = \hat{z}_t e^{(1-\sigma)g\epsilon}$ . This lets us rearrange and get:

$$\rho\tilde{N} = \pi + \lim_{\epsilon \rightarrow 0} \frac{e^{-\rho\epsilon}\tilde{N}(e^{(1-\sigma)g\epsilon} - 1)}{\epsilon}$$

This simplifies to:

$$\rho\tilde{N} = \pi - g(\sigma - 1)N$$

And recall that  $N = \tilde{N}\hat{z}$ , so  $N' = \tilde{N}$ .