NBER WORKING PAPER SERIES

TECHNOLOGICAL INNOVATION AND LABOR INCOME RISK

Leonid Kogan
Dimitris Papanikolaou
Lawrence D. W. Schmidt
Jae Song

Working Paper 26964
http://www.nber.org/papers/w26964

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2020

The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium. The authors also gratefully acknowledge financial support from the Becker-Friedman Institute for Research in Economics and the MIT Sloan School of Management. We thank Ufuk Akcigit, Effi Benmelech, Nicholas Bloom, Stephane Bonhomme, Jack Favilukis, Carola Frydman, David Hémous, Gregor Jarosch, Thibaut Lamadon, Ilse Lindenlaub, Elena Manresa, Adrien Matray, Derek Neal, Serdar Ozkan, Monika Piazzesi, Luigi Pistaferri, Martin Schneider, Kelly Shue, Toni Whited, and seminar participants in the Social Security Administration, Bank of Portugal, CEAR/GSU, Duke University, Duke/UNC Asset Pricing Conference, University of Chicago, UNC, the LAEF/Advances in Macro Finance workshop, Minnesota Asset Pricing Conference, MIT Sloan, University of Michigan, Northwestern, Stanford, SITE, and Yale for helpful comments and discussions. We are particularly grateful to Pat Jonas, Kelly Salzmann, and Gerald Ray at the Social Security Administration for their feedback during our SSA seminar presentation, help, and support. Further, we thank Josh Rauh, Irina Stefanescu, Jan Bena and Elena Simintzi for sharing their data and Fatih Guvenen and Serdar Ozkan for sharing their replication code. Brandon Dixon, Alejandro Hoyos Suarez, Kyle Kost, Andrei Nagornyi, Bryan Seegmiller, and Jiaheng Yu provided outstanding research assistance. The views expressed herein are those of the authors and not necessarily those of the Social Security Administration or the National Bureau of Economic Research.

Technological Innovation and Labor Income Risk
Leonid Kogan, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Jae Song
NBER Working Paper No. 26964
April 2020
JEL No. E24,G10,G12

ABSTRACT

We examine the relation between technological progress and the riskiness of labor income. Motivated by a simple model of creative destruction, we draw a distinction between technological innovation advanced by the firm, or its competitors. Using administrative data from the United States, we find that own firm innovation is associated with a modest increase in worker earnings growth, while innovation by competing firms is related to lower future worker earnings. Importantly, these earnings changes are asymmetrically distributed across workers: both gains and losses are concentrated on a subset of workers, which implies that the distribution of worker earnings growth rates becomes more right- or left-skewed following innovation by the firm, or its competitors, respectively. These effects are particularly strong for the highest-paid workers. Our results therefore suggest innovation is associated with a substantial increase in the labor income risk, especially for workers at the top of the earnings distribution. Our simulations reveal that the increased disparity in innovation outcomes across firms in the 1990s can account for a significant part of the recent rise in income inequality.

Leonid Kogan
MIT Sloan School of Management
100 Main Street, E62-636
Cambridge, MA 02142
and NBER
lkogan@mit.edu

Dimitris Papanikolaou
Kellogg School of Management
Northwestern University
2211 Campus Drive, Office 4319
Evanston, IL 60208
and NBER
d-papanikolaou@kellogg.northwestern.edu

Lawrence D. W. Schmidt
Sloan School of Management
Massachusetts Institute of Technology
100 Main Street
Cambridge, MA
ldws@mit.edu

Jae Song
Social Security Administration
Office of Disability Adjudication
and Review
5107 Leesburg Pike, Suite 1400
Falls Church, VA 22041
jae.song@ssa.gov

An appendix is available at: https://www.nber.org/data-appendix/w26964/

Income inequality in the United States has increased sharply over the last thirty years, reaching levels unprecedented during the post-war era. While the data indicate that firms have likely played a central role in this increase, the exact underlying mechanism remains unclear.[1] In this paper, we explore the extent to which differences in the rate at which firms acquire new technologies contribute to both workers' labor income risk and income inequality. We find that higher rates of innovation are associated with greater labor income risk—especially for workers at the top of the earnings distribution. In addition, we find that the increased disparity in innovation outcomes across firms in the 1990s has likely been an important factor in the rise of income inequality across firms.

To guide the empirical work, we start with a simple model of endogenous growth, in which firms are collections of product lines. Firms innovate randomly, at potentially different rates. Production of each good requires a worker/manager. Importantly, these workers' human capital is partly specific to a particular good/firm combination and incentive considerations imply that workers' incomes are sensitive to firm profits. Firm innovation can take the form of an improvement in one of its own product lines, or by acquiring the technological lead in a product line owned by another firm. If the firm innovates on its own good, the existing worker either receives a wage increase, or is replaced, depending on whether the improvement takes the form of an improvement in the quality of the product or an improvement in production methods. If the firm loses a product line to its competitors, the worker is displaced.

An increase in the rate of innovation by the firm can therefore lead to an increase in both the mean, and the dispersion, in the path of future earnings for the firms' incumbent workers. While some workers are better off since their specific human capital becomes more productive, others are replaced and experience subsequent income declines. By contrast, increases in the rate of innovation by competing firms are associated with more negatively skewed earnings growth. A subset of employees are displaced and experience large income declines, whereas others are largely unaffected. Innovation is thus linked with an increase in the risk of human capital; highest paid workers face more downside risk because they have the most to lose if their specific human capital is displaced—since their outside option is not proportional to their prior wage.

The model is simple, yet it leads to rich and testable predictions about links between innovation and the distribution of worker earnings. First, innovation by the same firm is associated with an increase in the dispersion of worker earnings for incumbent workers. Second, innovation by competing firms leads to more negatively skewed earnings growth. Third, both of these effects are larger in magnitude for higher-paid workers. Fourth, the imputed profit-sharing elasticities—comparing the change in worker earnings to change in profits—is larger in response to competitor innovation. This effect arises because the possibility that even own firm innovation can be associated with displacement of a worker's specific human capital creates a wedge between expected firm and worker-level outcomes that is not present for competitor innovation.

[1] See, for instance Song, Price, Guvenen, Bloom, and Von Wachter (2019); Barth, Bryson, Davis, and Freeman (2016).

To empirically examine these predictions, we combine administrative data on worker earnings from the Social Security Administration (SSA) with the firm-level measure of the value of innovation developed in Kogan, Papanikolaou, Seru, and Stoffman (2017). Kogan et al. (2017) propose a measure of the economic importance of each innovation that combines patent data with the stock market response to news about these patents. An advantage of this measure is that it allows us to connect each new invention or production method to its originating firm, and therefore isolate innovation by the worker's own firm from innovation by its competitors. Kogan et al. (2017) show that their measure is strongly related to changes in ex-post profitability across firms and document evidence consistent with creative destruction. We find similar reallocative effects for worker earnings. Indeed, we find that increases in firm innovation are associated with increased future earnings growth for the firms' own workers. The magnitude of these effects is sizeable and own-firm effects are consistent with most extant estimates of profit sharing elasticities. Importantly, we also document that more valuable innovation by competing firms is associated with significantly lower future worker earnings. Moreover, holding fixed the associated change in expected firm profits, workers' earnings growth is more sensitive to innovation by competing firms relative to own firm innovation.

Importantly, and consistent with the theoretical model, these effects are not distributed symmetrically across all workers in the same firm. In particular, we use quantile regressions to characterize how the entire distribution of worker earnings growth rates shifts following innovation by the firm, or its competitors. We find that subsequent to innovation by their own firm, the distribution of earnings for the firm's own workers becomes more positively skewed. That is, the increase in average earnings we documented above is concentrated among a small subset of workers. Conversely, innovation by competitors is associated with more negatively skewed earnings growth for the firm's workers. Specifically, most workers experience small declines in income, while a minority experiences a significant drop in its labor earnings.

Thus, in addition to changes in the conditional mean of worker earnings growth, innovation by either the firm or its competitors is empirically associated with higher labor income risk—that is, shifts in the higher moments of income growth. This raises the question of whether this increase in ex-post heterogeneity in worker outcomes indeed corresponds to increased income risk, or whether it is related to ex-ante worker characteristics. For instance, it could be that all innovations are complementary to high-skill workers, and since we do not observe worker skill, the increase in dispersion in earnings growth rates in response to innovation that we are picking up could simply be an increase in the skill premium. Though worker skill is unobservable, it is likely correlated with income levels. Therefore, we next allow our estimates to vary depending on the worker's earnings rank within the firm (net of life-cycle effects) as a proxy for her skill level. We find that conditioning on earnings levels does little to reduce the degree of heterogeneity in worker outcomes, which is consistent with the idea of higher uncertainty in future income.

Importantly, we find that the sensitivity of earnings growth to innovation outcomes (either by the firm, or its competitors) is greater for the highest paid workers (top 5% within the firm). Specifically, the increase in the positive skewness of income growth in response to firm innovation is higher for its top workers; these top workers however also experience an increase in the left tail of earnings growth when their own firm innovates. Top workers are also significantly more adversely affected than the average worker to innovation by competing firms: top workers experience a 14 percentage point reduction in the 5th-percentile of earnings growth in response to a one-standard deviation increase in the average level of innovation by competing firms—compared to 2.1 to 6 percentage points reduction for other workers. These findings are particularly striking in light of the traditional view that technology tends to complement high-skill labor (Goldin and Katz, 2008).

In brief, our quantile regression estimates reveal that higher innovation is associated with significant increases in the dispersion of labor income. In the model, earnings losses are driven by job loss. An advantage of our data is that they allow us to track workers across firms, and therefore examine the extent to which this increase in the risk of earnings declines is related to separations. Consistent with our model, we find that workers employed in firms that do not innovate while their competitors do are more likely to subsequently exit the firm. Since the distribution of earnings growth rates is substantially more negatively skewed for exiting relative to continuing workers, a higher chance of job loss following high innovation outcomes by competing firms partially accounts for the higher negative skewness in earnings growth rates.

However, we also find a correlation between the rate of innovation and shifts in the distribution of earnings growth even among the subset of exiting workers. Specifically, we find that workers that leave the firm following periods of high innovation, either by the firm or its competitors, are significantly more likely to experience large subsequent declines in their labor income than workers which leave during periods of low innovation. We find that extended periods of unemployment (years of zero earnings) following periods of high innovation are partly responsible for this increase. Workers that exit following periods of high innovation are also more likely to apply for disability insurance than workers exiting following periods of low innovation. These patterns go beyond our simple model, but they suggest that innovation can have a persistent effect on worker productivity that extends beyond job loss.

We conclude our analysis with a decomposition exercise which quantifies the role of innovation for the recent rise in income inequality among firms. A stylized fact in the data is that the aggregate amount of innovation increased during the 1990s. Importantly, this increase in the level was also accompanied by an increase in the dispersion of innovation outcomes across firms (even within industries). That is, most of the increase in the amount of innovation was concentrated among a relatively small subset of firms. By simulating from our estimated quantile regression model, we show that this increase in the dispersion in firm innovation outcomes can account for much of the increase in between-firm inequality during the last few decades. In terms of within-firm inequality, we find both the increase in the level as well as the dispersion in innovation play a role.

An important caveat in our analysis is that the statistical relations we document need not be causal. For instance, workers in R&D-intensive firms may have a different earnings structure than workers in other firms. Though we cannot exclude the possibility that omitted variables are the main drivers of our results, several factors mitigate this concern. First, our innovation measures are strongly related to future firm profits, but are uncorrelated with past trends in firm profitability. Second, the response of the distribution of worker earnings to innovation is qualitatively distinct from its response to changes in profitability or stock returns—particularly in regards to competitor outcomes. This difference suggests that the effects we are picking up are specific to innovation outcomes per se, as opposed to shifts in underlying profitability trends at the industry level. Third, our point estimates are essentially unchanged if we expand the set of covariates to include controls for past R&D spending. In this case, we are comparing firms that spent the same resources on R&D, and exploiting the fact that some firms produce patents that generate a larger stock market reaction than other firms. Last, the fact that we are measuring patent values based on stock market reactions—which should be unexpected—mitigates the issue, though only on the intensive margin.

Our work contributes to our understanding of the role played by technological innovation and displacement in the product market on the distribution of worker earnings—both in terms of shifts in labor income risk as well as changes in income inequality. Our focus on firms is motivated by the well-documented large and persistent differences in firm productivity (Syverson, 2011); the importance of firms in understanding worker earnings inequality (Song et al., 2019); and the evidence that workers share in firm profits (see, e.g., Card, Cardoso, Heining, and Kline, 2018, for a survey).

Our focus on firms distinguishes our work from most of the existing work studying the link between technological innovation and worker earnings. That is, existing work has emphasized the complementarity between technology and certain types of worker skills (Goldin and Katz, 1998, 2008; Autor, Levy, and Murnane, 2003; Autor, Katz, and Kearney, 2006; Goos and Manning, 2007; Autor and Dorn, 2013); or the substitution between workers and new forms of capital (Hornstein, Krusell, and Violante, 2005, 2007; Acemoglu and Restrepo, Acemoglu and Restrepo). Most of these papers do not distinguish between innovation and adoption. By contrast, our use of patent data implies that we necessarily focus on innovation rather than adoption of existing technologies. As a result, we draw a distinction among innovations depending on which firm they originate in. In this regard, our work is closest to the literature that studies the impact of firm innovation on the earnings of its own workers (van Reenen, 1996; Aghion, Bergeaud, Blundell, and Griffith, 2017; Kline, Petkova, Williams, and Zidar, 2019; Howell and Brown, 2020). The central finding in this body of work is that innovative firms pay higher wages to incumbent workers, consistent with ex-post sharing of quasi-rents. We view our work as complementary; rather than focusing on estimating rent-sharing elasticities, our main goal is to understand the relation between technological progress, product market competition, and the entire distribution of worker earnings changes—shifts in labor income risk and earnings inequality. Our work is also related to Aghion, Akcigit, Bergeaud, Blundell, and

Hemous (2018), who document a correlation between regional differences in patenting activity and top income inequality.

Our work is related to the literature arguing for the importance of firms for understanding the dynamics of income inequality. Abowd, Kramarz, and Margolis (1999) propose that firm heterogeneity accounts for a substantial fraction of wage differences across workers. Song et al. (2019) document that a substantial fraction of the rise in income inequality across workers can be attributed to increasing differences in average worker pay across firms. Card, Heining, and Kline (2013) find similar effects in Germany. For firms to play a role in inequality, they need to share rents with workers. Card et al. (2018) survey the literature on estimating rent-sharing elasticities between workers and firms; most recent studies that employ micro data deliver estimates that lie between 0.05 to 0.15. For our purposes, the most directly relevant estimates are those of Lamadon, Mogstad, and Setzler (2019), who estimate an coefficient of 0.13-0.14 using recent IRS data from the US. Our OLS point estimates for stayers that compare the increase in profitability to the increase in the earnings of the average worker following innovations by the firm are somewhat higher than this range (0.195), but are closer to the estimates reported in van Reenen (1996) and Kline et al. (2019), who report elasticities of 0.29 and 0.19–0.23, respectively. In addition, our quantile regressions reveal substantial heterogeneity in worker outcomes following technological improvements by the firm—or its competitors—that are otherwise obscured. This is particularly important, in light of the fact that the existing literature has often interpreted these elasticities as a measure of the degree of insurance provided by the firm's owners to workers (Guiso, Pistaferri, and Schivardi, 2005; Lagakos and Ordoñez, 2011; Fagereng, Guiso, and Pistaferri, 2018). Our findings illustrate that focusing on average responses can mask substantial heterogeneity in ex-post outcomes across workers.

Last, our work has important implications for the literature studying the asset pricing implications of skewness in labor income. For example, Constantinides and Ghosh (2017) and Schmidt (2016), argue for the importance of 'income tail risk' for the equity premium. Their estimates build on the counter-cyclical nature of the left-skewness of idiosyncratic income changes documented by Guvenen, Ozkan, and Song (2014). In terms of magnitudes, the effects we document are comparable: focusing on top workers, the increase in the left tail of income growth following periods of innovation by competing firms in the same industry is comparable in magnitude to the increase in the left tail documented in recessions documented in Guvenen et al. (2014). Our findings, therefore, directly relate to the recent work arguing that investors might want to purchase insurance against states with high degrees of technological innovation (Papanikolaou, 2011; Garleanu, Kogan, and Panageas, 2012; Kogan, Papanikolaou, and Stoffman, 2020).

1 Theoretical Framework

We begin by proposing a simple model that serves to guide the empirical work. The model is based on a relatively standard quality-ladder model in continuous time in the spirit of Aghion and Howitt

(1992), modified to include human capital.

Aggregate output is produced by a continuum of intermediate goods $x_{i,t}$

$$X_t = \int_0^1 x_{i,t}^{\nu} di, \qquad \nu \leq 1. \tag{1}$$

Each intermediate good can be produced by a continuum of competitive firms, which however differ in their level of efficiency in producing each good. Firms produce each good using a constant returns to scale technology; hence, we can assume that only the most productive firm finds it profitable to produce each good. In particular, good i is produced using the following technology,

$$x_{i,t} = q_{i,t}\, e_{i,t}\, l_{i,t}, \tag{2}$$

where $q_{i,t}$ is the quality of the leading producer and $l_{i,t}$ is a second factor of production (unskilled labor, robots, or land) that can be freely reallocated across product lines. In addition, the production of an intermediate good requires a skilled worker or a manager. There is a moral hazard friction, in that the skilled worker can potentially divert output—hence, $e_{i,t}$ denotes the fraction of un-diverted output. However, diversion is costly: if she diverts 1 unit of output, she can only effectively steal a fraction β. Hence, the (static) solution to this moral hazard friction is to provide the manager with a fraction β of the profits from producing good i, in which case she is indifferent between stealing versus not. In what follows, we assume that this is the case, which implies that there is no output diversion in equilibrium—and hence $e_{i,t} = 1$.

Given our assumptions, the total profits from producing good i—to be shared by the skilled worker and the firm's owners—are equal to

$$\Pi_{i,t} = A_t\, q_{i,t}^{\frac{\nu}{1-\nu}}, \tag{3}$$

where

$$A_t \equiv \nu \left(1 - \frac{1}{\kappa}\right) \left[\int_0^1 q_{i,t}^{\frac{\nu}{1-\nu}}\, di\right]^{-\nu} \tag{4}$$

depends on the distribution of leading quality $q_{i,t}$ across goods.

A firm is a collection of goods that finds it profitable to produce—that is, goods in which the firm is the leading producer. Innovation is exogenous and takes the form of improvements in efficiency: over an instant dt, a firm can innovate with probability $\lambda_{f,t}\, dt$. Here, $\lambda_{f,t} \in (\lambda_L, \lambda_H)$ is a two-point Markov process, with transition probability from state s to s' given by $\mu_{ss'}\, dt$. Conditional on successfully innovating, a firm either improves upon one of the goods it is already producing (with probability μ) or (with probability $1 - \mu$) on one of the goods produced by another firm f' in the set $C(f)$—that is, one of the firm's 'competitors'. For illustrative purposes, assume each firm f has one competitor f'. Quality improvements are proportional, so conditional on innovation, the firm's production efficiency $q_{i,t}$ increases by a factor $\kappa > 1$.

On the firm side, the model we have described so far is relatively standard. An increase in the rate at which the firm innovates $\lambda_{f,t}$ leads to an increase in firm profits and employment, as we can see in Panel A of Figure 1. Profits increase more than employment because innovation partly consists of improving existing varieties, which do not require the firm to hire additional workers. By contrast, increases in the rate of innovation by the firm's competitors (here, firm f') is associated with lower profits and employment, since the likelihood that the firm loses the technology lead in one of its own products increases.

However, innovation can potentially displace workers. Specifically, if a firm improves upon one of the goods it is already producing, the skilled worker is retained with probability p; with probability $1-p$ the firm hires a new worker (from the unemployed pool). One interpretation here is that the firm can improve on a good either by improving the quality of the product, or by improving its production method. In the former case, the worker is retained; in the latter case, the worker is potentially replaced. In addition, once a firm f loses the leading efficiency to a competitor f', the position is eliminated and the worker previously assigned to that good becomes unemployed; firm f' hires a new worker from the unemployed pool. Unemployed workers receive a benefit b_t, which is equal to a fraction of the salary of the least-paid manager.

The goal of the model is to illustrate the implications of firm innovation for the distribution of worker earnings growth. Due to the moral hazard friction, the skilled worker j assigned to good i receives a fraction of the firm's profits,

$$w_{j,t} = \beta\, A_t\, q_{i,t}^{\frac{\nu}{1-\nu}}. \tag{5}$$

Consider the earnings at time $t + dt$ for the worker j that is currently assigned to good i (owned by firm f) at time t. Over the next interval dt, her earnings growth evolves according to,

$$\log\left(\frac{w_{j,t+dt}}{w_{j,t}}\right) - d\log A_t = \begin{cases} \dfrac{\nu}{1-\nu}\log\kappa > 0 & \text{with prob. } p\,\mu\,\lambda_{f,t}\,dt \\ \log\left(\dfrac{b_t}{w_{j,t}}\right) < 0 & \text{with prob. } \left[(1-\mu)\,\lambda_{f',t} + (1-p)\,\mu\,\lambda_{f,t}\right]\,dt. \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Equation (6) illustrates how a shock to the rate of innovation by a firm $\lambda_{f,t}$, or its competitor $\lambda_{f',t}$, is likely to affect the future distribution of earnings growth for a given worker. A positive shock to $\lambda_{f,t}$ increases the likelihood of both an increase as well as a decrease in worker earnings. By contrast, an increase in the rate of innovation by competing firms $\lambda_{f',t}$ increases the likelihood of large earnings losses. Importantly, the magnitude of earnings losses is increasing in the worker's current wage $w_{j,t}$ due to the assumption that her continuation value after displacement (which is equal to b_t) is independent of her current wage.

Panel B of Figure 1 illustrates these effects by presenting the sensitivity of different percentiles of worker earnings growth over the next five years to an increase to the rate of innovation by a firm $\lambda_{f,t}$ or its competitors $\lambda_{f',t}$. We see that a positive shock to the rate of competitor innovation $\lambda_{f',t}$ lowers the mean wage growth of incumbent workers by making the distribution of earnings growth more negatively skewed. By contrast, a positive shock to firm innovation $\lambda_{f,t}$ leads to a moderate increase in both the average as well as the dispersion in earnings growth rates—that is, it is associated with both a mean shift, but also a widening of the worker earnings distribution. In both cases, the downside exposures of higher income workers are significantly larger, whereas outcomes look fairly similar at the median and in the right tail.

The key model mechanism that drives these findings is profit sharing—due to the moral hazard friction—combined with the specificity of a worker's human capital. Specifically, profit sharing implies a pass-through from firm to worker earnings. Human capital specificity to a particular firm/product combination implies that innovation is associated with the likelihood of job loss. Naturally, this specificity encompasses several mechanisms through which technological innovation may affect labor income—such as automation of certain tasks or skill displacement. Importantly, we view the 'managers' in our model as skilled workers, with some of these skills being specific to a particular firm or production method. By contrast, we can interpret the second, flexible factor $l_{i,t}$ as including unskilled workers, who are likely to be equally productive across all production methods or firms.

Last, the model implies that the 'profit-sharing' elasticity, that is, the sensitivity of worker earnings to firm profits varies with the nature of the shock. By comparing the mean increase in worker earnings in Panel B to the increase in firm profits in Panel A, we can see that the ratio of the former to the latter is much greater in response to a shock to competitor innovation $\lambda_{f',t}$ than in response to innovation by the firm $\lambda_{f,t}$. The reason is that a positive shock to $\lambda_{f',t}$ has an unambiguous adverse effect on both firm and worker earnings. By contrast, a positive shock to $\lambda_{f',t}$ increases firm profits, but can lead to earnings declines for some workers. In both cases, costs are concentrated on a subset of workers whose human capital is displaced, so innovation is accompanied by an increase in earnings risk, especially for the highest paid workers.

In sum, the simple model we have outlined in this section makes the following predictions. First, innovation by the same firm is associated with an increase in the dispersion of worker earnings for incumbent workers. Second, innovation by competing firms leads to more negatively skewed earnings growth. Third, both of these effects are larger in magnitude for higher-paid workers. Fourth, the imputed profit-sharing elasticities—comparing the change in worker earnings to change in profits—is larger in response to competitor innovation. The main part of the paper focuses in analyzing these predictions in the data.

2 Data and Measurement

Here, we briefly summarize the data on labor income and firm innovation outcomes used in our analysis. All details are relegated to Appendix B.

2.1 Labor Income

Our data on worker earnings are based on a random sample of individual records for males, drawn from the U.S. Social Security Administration's (SSA) Master Earnings File (MEF). The MEF includes annual earnings information for every individual that has ever been issued a Social Security Number. The earnings data are based on box 1 of the W2 form, which includes wages and salaries; bonuses; the dollar value of exercised stock options and restricted stock units; and severance pay. The data are based on information that employers submit to the SSA, and are uncapped after 1978. Importantly, the data have a panel structure, which allows us to track individuals over time and across firms.

Our main sample covers the 1980–2013 period, and the data construction closely follows Guvenen et al. (2014).[2] In addition, we closely follow Guvenen et al. (2014) and impose several additional filters to the data that exclude self-employed workers and individuals with earnings below a minimum threshold, which is equal to the amount one would earn working 20 hours per week for 13 weeks at the federal minimum wage. See Appendix B.1 and Guvenen et al. (2014) for further details.

Our key outcome variables of interest are growth rates of income, cumulated over various periods, and adjusted for life cycle effects. We follow closely Autor, Dorn, Hanson, and Song (2014) and construct a measure of a worker's average earnings between periods t and $t+k$, that is adjusted for life-cycle effects:

$$w^i_{t,t+h} \equiv \log\left(\frac{\sum_{j=0}^{h} \text{W-2 earnings}_{i,t+j}}{\sum_{j=0}^{k} D(\text{age}_{i,t+j})}\right). \tag{7}$$

Here, W2 earnings$_{i,t}$ is the sum of earnings across all W-2 documents for person i in year t. In the denominator, $D(\text{age}_{i,t})$ is an adjustment for the average life-cycle path in worker earnings that closely follows Guvenen et al. (2014). In the absence of age effects, $D(\text{age}_{i,t}) = 1$, hence (7) can be interpreted as (the logarithm of) the average income from period t to $t+h$, scaled by the average income of a worker of a similar age.

Equation (7) describes a worker's *age-adjusted earnings*; to conserve space, we will simply refer to it as worker earnings. When focusing on worker earnings growth, our main variable of interest will be the cumulative growth in (7) over a horizon of 5 years:

$$g_{i,t:t+5} \equiv w^i_{t+1,t+5} - w^i_{t-2,t}. \tag{8}$$

[2]Specifically, a sample of 10 percent of US males are randomly selected based on their social security number (SSN) in 1978. For each subsequent year, new individuals are added to account for the newly issued SSNs; those individuals who are deceased are removed from that year forward. We start our analysis in 1980 to overcome potential measurement issues in the initial years following the transition to uncapped earnings.

Examining (8), we note that the base income level over which growth rates are computed is the average (age-adjusted) earnings between $t-2$ and t. Focusing on the growth of average income over multiple horizons in (8) emphasizes persistent earnings changes, and therefore helps smooth over large changes in earnings that may be induced by large transitory shocks or temporary unemployment spells (see Appendix B.3 for more details). In our baseline case, we will consider the ratio of 5-year forward earnings to the last 3 years of cumulative earnings (note that we simulated the same quantity in the model above). Altering the forward window allows us to explore the persistence of our findings. For brevity, we restrict attention to a backward window of 3 years.

A key advantage of our data is that they allow us to track employees across firms. Therefore, when computing (8) for a given worker i at time t, her earnings growth rate may include income from more than one firm if she were to switch employees at some point between years $t+1$ and $t+5$. In some cases, we will distinguish earnings growth for workers that move or remain with the current employer—examine mobility as a separate outcome. Since we want to allow for delayed effects on mobility, but at the same time capture worker earnings changes in the new job, we will define a mover as those workers that do not work in the same firm at $t+3$ as they did in year t. Consequently, the earnings growth of a mover will include the change in her salary from moving out of the current firm. Stayers are defined as workers who did not move between t and $t+3$.

2.2 Innovation Outcomes

Our main independent variables of interest are innovation outcomes at the firm level. The most broadly available data on innovation are based on patents. An advantage of using the patent data is that they can be linked to the firm level, which allows us to separately estimate the relation between worker earnings and innovation by the firm and its competitors. Hence, importantly, our definition of 'innovation' will be somewhat narrow as a result. That is, we will *not* be measuring firms' adoption of technologies developed by other firms. Hence, our results will be rather distinct from the literature focusing on the complementarity between skilled workers and new types of capital goods (for example, robots, as in Acemoglu and Restrepo (Acemoglu and Restrepo)).

We first construct an empirical analogue of $\lambda_{f,t}$ and $\lambda_{f',t}$ in the model in Section 1. In the model, all innovations are equal in quality, since κ is constant. In the data, they are likely not. Indeed, a major challenge in measuring innovation by using patents is that they vary greatly in their technical and economic significance (see, e.g., Hall, Jaffe, and Trajtenberg, 2005; Kogan et al., 2017). We will, therefore, be weighting individual patents by their estimated market value using the data by Kogan et al. (2017), henceforth KPSS, who develop an estimate of the market value of a patent based on the fluctuations in the stock price of innovating firms following patent grants. Thus, their measure is only available for public firms. We, henceforth, refer to their measure as the 'market' value of a patent.

We follow KPSS closely and construct measures of the value of innovation by the firm

$$A_{f,t} = \frac{\sum_{j \in P_{f,t}} \xi_j}{B_{ft}} \tag{9}$$

and its competitors,

$$A_{I \setminus f,t} = \frac{\sum_{f' \in I \setminus f} \left(\sum_{j \in P_{f',t}} \xi_j \right)}{\sum_{f' \in I \setminus f} B_{f't}}. \tag{10}$$

Here, ξ_j corresponds to the KPSS value of a patent. The set of competing firms $I \setminus f$ is the 'leave-out mean'—defined as all firms in the same SIC3 industry, excluding firm f. Large firms tend to file more patents. As a result, both measures of innovation above are strongly increasing in firm size (Kogan et al., 2017). To ensure that fluctuations in size are not driving the variation in innovative output, we follow KPSS and scale the measures above by firm size. We use book assets as our baseline case, but our main results are similar if we scale by the firm's market capitalization instead (see, e.g., Appendix Figure A.16). Appendix B.4 provides more details on the construction of these variables. In the context of the model in Section 1, we can interpret $A_{f,t}$ and $A_{I \setminus f,t}$ as empirical proxies of $\lambda_{f,t}$ and $\lambda_{f',t}$.[3]

A potential shortcoming of patent-based measures of innovation is that the exact timing of its impact on firm wages is somewhat ambiguous. A successful patent application helps the firm appropriate any monopoly rents associated with that invention, hence dating patents based on their issue date seems like a natural choice. The patent issue date is also the date at which the fact that the patent application is successful, and therefore forms the basis for estimating the value of the patent based on the firm's stock market reaction in KPSS. For our purposes, however, this timing choice may be somewhat problematic when examining how worker earnings respond to the firm's own innovation. For instance, the firm may decide to pay workers in advance of the patent grant date. Hence, income changes subsequent to the patent grant date may be affected by temporary increases in worker salaries prior to the patent grant date. To address this concern, we date the firm's own patents based on the year that these patents are actually filed. Hence, when computing $A_{f,t}$, the set of patents $P_{f,t}$ includes patents that are *filed* in year t.[4] Consistent with this timing convention, Appendix Figure A.3 indicates that firm profits respond sharply in the year immediately after patents are filed, despite the fact that most patents take several years to be approved, and are

[3] In particular, many firms have hundreds or thousands of patent applications in a given year. Many of these innovations, however, are likely to be incremental. Weighting by the estimate of the market value of a patent helps down-weigh more marginal patents, but the result is still a continuous measure which is likely to be a noisy estimate of the underlying level of firm innovation. We, therefore, interpret a high value of $A_{f,t}$ as indicative of a higher likelihood that the firm has improved its efficiency in a given product—that is, as a positive shock to $\lambda_{f,t}$. An alternative strategy would have been to only focus on patents on the right tail of the distribution of $A_{f,t}$; however, doing so would require us to impose an arbitrary threshold.

[4] Patent applications (and hence, filing dates) are only disclosed ex-post. Hence, the value ξ_j is still computed using the market reaction on the patent grant date. Our implicit assumption is that this value represents a known quantity to the firm as of the application date, similar to the assumptions regarding the number of future citations a patent receives that are common in the innovation literature (see, e.g., Hall et al., 2005).

associated with substantially larger cumulative responses of profits. Patents, by competing firms used in the construction of $A_{I\setminus f,t}$, are dated as of their issue date. That said, this choice of timing is not a main driver of our findings on earnings growth rates, as most results are qualitatively similar if we date the firm's patents as of their grant date.

2.3 Overview of the sample

Our final matched sample includes approximately 14.6 million worker-years observations. Appendix Table A.1 provides some summary statistics for the key variables in our analysis. To arrive at this sample, we merge the firm-level data on innovation with individual workers' earnings histories using EIN numbers. On average, matching rates are quite high: we can find records in the MEF for about 84% of the public firm-years (see Appendix Tables A.2 to A.3 and Figure A.1 for further details). The industry composition of the matched and unmatched sample is similar. Matched firms tend to have similar levels of book assets and somewhat higher levels of employment (as reported on 10-K forms) and innovative activity than the unmatched sample of public firms. In terms of the workforce composition, employees at matched public firms are slightly older; earn about \$16 thousand dollars more per year; and have worked on average slightly longer in the same firm.

3 Innovation, profits, and worker earnings

Here, we examine the link between firm innovation, firm profits, and worker earnings. The simple model we outlined in Section 1 helps guide the empirical work. In particular, the model predicts that firm innovation leads to a more right-skewed distribution for worker earnings growth. By contrast, the distribution of worker earnings should become more left-skewed in response to innovation by competing firms. Last, the model implies that the increase in the left tail should be greatest for the firm's top workers. In what follows, we will examine these predictions in more detail.

3.1 Firm Innovation

We begin by describing the behavior of the firm innovation measure A_f during our sample period in Figure 2. Panels A and B plot the average level of innovation during our period. We see that the 1990s was a particularly innovative period—which is in line with the evidence in KPSS, among others. Panel C plots the dispersion in the A_f measure across firms; we see that the cross-sectional dispersion in firm innovation outcomes (as measured by the coefficient of variation in A_f) exhibited a similar increase as the level. Panel D plots the share of aggregate innovation that can be attributed to the most innovative firms in each year (the top 1% innovation share). We see that this period was also associated with an increase in the concentration of innovative outcomes: during the late 1990s, the top 1% of innovative firms accounted for over 50% of the total innovative output in our sample, compared to 35% in the mid 1980s.

One possibility is that these facts may simply be indicative of differential industry trends during this period. However, this is not the case. Specifically, Panels E and F decompose the dispersion in innovation outcomes A_f across firms into a within- and a between-industry (SIC3) measure. As we see in Panels E and F, this increase in dispersion is primarily a within-industry phenomenon. By contrast, we find no comparable shifts in the dispersion of outcomes across industries.

In brief, the data suggest that the innovation boom in the 1990s was primarily driven by a small set of firms in each industry. The natural question is then what do these facts imply for the distribution of worker earnings. In the context of the model in Section 1, we can interpret these shifts as an increase in the dispersion of $\lambda_{f,t}$ across firms. We will revisit these findings in Section 6 and explore their implications for inequality in worker earnings.

3.2 Innovation and Profitability

Kogan et al. (2017) show that differences in innovation outcomes are associated with substantial heterogeneity in subsequent growth in profitability and employment, and they document sizeable creative destruction effects. To set the stage for what follows, we begin by revisiting part of their analysis. To make our results comparable with our worker-level regressions, we estimate a slightly modified specification than KPSS that closely parallels our worker earnings growth measure (8). That is, we estimate

$$\log\left[\frac{1}{|h|}\sum_{\tau=1}^{h} X_{f,t+\tau}\right] - \log X_{f,t} = a_h\, A^{sm}_{f,t} + b_h\, A^{sm}_{I\backslash f,t} + c_h\, Z_{ft} + u_{ft+h}, \tag{11}$$

the dependent variable in equation (11) is the growth in the average level of profits and employment over the next h years. Thus, equation (11) can be interpreted as the analogue of the model impulse responses in Figure 1 in the data.

The vector Z includes several controls, including one lagged value of the dependent variable and the log of the book value of firm assets to alleviate our concern that firm size may introduce some mechanical correlation between the dependent variable and our innovation measure. For instance, large firms tend to innovate more, yet grow slower (see, e.g., Evans, 1987); controlling for other measures of size yields similar results. We also control for firm idiosyncratic volatility σ_{ft} because it may have a mechanical effect on our innovation measure and is likely correlated with firms' future growth opportunities or the risk in worker earnings. Further, we include industry and time dummies to account for unobservable factors at the industry and year level. We cluster standard errors by firm and year. To evaluate economic magnitudes, we normalize A_f and $A_{I\backslash f}$ to unit standard deviation. Last, we restrict the sample to the time period of the SSA data (1980–2013), though similar estimates obtain in the full sample.

Our main coefficients of interest are a_h and b_h, which measure the response in firm profits, employment, and productivity to innovation by the firm and its competitors, respectively. Panel

A of Table 1 presents our estimates for horizons h of 3, 5, and 10 years. We see that future firm profitability is strongly related to the firm's own innovative output. The magnitudes are substantial; for instance, a one-standard deviation increase in firm's innovation is associated with an increase of approximately 8% in the average level of profits over the next 5 years. Similar to KPSS, the estimates of b suggest that innovation is associated with a substantial degree of creative destruction. In particular, a one-standard deviation increase in innovation by firm's competitors is associated with a decline of 4.9% in the level of profits over the next 5 years. As we compare the estimates between horizons of five to ten years, we see that these are largely permanent effects.

In sum, we see that own-firm innovation is associated with increased firm profitability. By contrast, firms that do not innovate when their competitors do, experience declines in profits. These results are consistent with models of endogenous growth, in which firm innovation is an important driver of profitability. However, another possibility is that it is related to some unobservable source of heterogeneity that itself is responsible for increased firm profits. Hence, we next examine whether innovation is related to past trends in profitability. That is, we re-estimate equation (11), but now allow h to take negative values. Appendix Figure A.2 plots the estimated coefficients a_h and b_h for values of $h = -5$ to $h = 10$. Examining both panels of the figure, we see that the relation between innovation by the firm ($A^{sm}_{f,t}$) or its competitors ($A^{sm}_{I\setminus f,t}$) at time t and profitability prior to year t is essentially zero. These results reveal that our innovation measure is not related to pre-existing trends in profitability. In addition, we view these results as supportive of our convention for dating patents.[5]

3.3 Innovation and Mean Worker Earnings

Next, we examine what happens to worker earnings in response to innovation by the firm or its competitors using a similar specification as above

$$g_{i,t:t+h} = a_h\, A^{sm}_{f,t} + b_h\, A^{sm}_{I\setminus f,t} + c_h\, Z_{i,t} + \varepsilon_{i,t}, \tag{12}$$

where $g_{i,t:t+h}$ is the (cumulative) growth in employee i's labor income over the next h years — defined in equation (8) above, and A_f and $A_{I\setminus f}$ are the measures of firm and competitor innovation defined in equations (9) and (10) above. Our vector of controls Z includes a similar set of firm-level controls as our firm-level regressions above (11); in addition, we saturate our specification with a battery of controls that aim to soak up ex-ante worker heterogeneity. Specifically, we include flexible non-parametric controls for worker age and past worker earnings as well as recent earnings growth

[5] As a further robustness check, we also estimated equation (11) using alternative choices for the timing of innovation. Consistent with our prior, we see a somewhat larger response of firm-level outcomes to own firm innovation when we date patents according to their filing as opposed to their grant date. Conversely, the relation with competitor innovation is stronger when competitor patents are dated according to their issue date. See Appendix Figure A.3 for more details.

rates.[6] To ensure that our point estimates are comparable to the analysis above (in which the unit of observation is at the firm-year as opposed to the worker-year level), we weigh observations by the inverse of the number of workers in each firm-year. We compute standard errors using a block-resampling procedure that allows for persistence at the firm level (that is, the analogue of clustering by firm). See Appendix C.1 for more details.

Panel B of Table 1 reports the estimated coefficients a and b, which capture the relation between average future worker earnings and innovation by the worker's own firm, or its competitors, respectively. As before, we examine horizons of 3, 5 and 10 years. Focusing again on the 5-year horizon, we see that a one-standard deviation increase in A_f is associated with a cumulative increase of 1.4% to the average worker earnings in the firm. By contrast, a one-standard deviation increase in innovation by competing firms is followed by a 1.9% decline in average worker earnings in firms that do not innovate. Comparing across columns (horizons) in each panel, we again see these are associated with essentially permanent changes in worker earnings.

One way of assessing the economic magnitudes of these coefficients is by relating them to the findings of the literature on estimating profit-sharing elasticities. Specifically, we compare the estimated magnitude of the responses in average worker earnings to firm innovation to the response of firm profitability in Section 3.2 above. Focusing on the 5-year horizon, we see from Panels A and B that a one-standard deviation increase in A_f is associated with a 7.2% increase in profitability compared to a 1.4% increase in earnings for the firm's own workers. These numbers imply a profit-sharing elasticity approximately equal to $1.4/8.0 \approx 0.17$. To put this number in context to the literature, we compare it to van Reenen (1996) and Kline et al. (2019), since their setting is most comparable to ours. These studies report elasticities of 0.29 and 0.19 (all workers)–0.23 (firm stayers), respectively.

Importantly, however, we note that the profit sharing elasticity that is implied by examining the response to competitor innovation is much larger. Specifically, focusing now on the 5-year horizon, we see that a one-standard deviation increase in innovation by competing firms $A_{I \setminus f}$ is associated with a 4.9% decline in profitability and a 1.9% decrease in earnings for the firm's own workers—implying a rent-sharing elasticity of $1.9/4.9 \approx 0.38$. Thus, our estimates suggest that declines in profits associated with competitor innovation are passed through at a higher rate than the benefits from own firm innovation. This finding is consistent with the model: since firm innovation may lead to replacing a worker with a new one, the expected earnings growth of incumbent workers is smaller than the firm's increase in profits

In sum, we find that innovation is strongly related to firm growth and to creative destruction.

[6] We construct controls for worker age and lagged earnings $w^i_{t-4,t}$ by linearly interpolating between 3^{rd} degree Chebyshev polynomials in workers' lagged income quantiles within an industry-age bin at 10-year age intervals. In addition, to soak up some potential variation related to potential mean-reversion in earnings (which could be the case following large transitory shocks), we also include 3rd degree Chebyshev polynomials in workers' lagged income growth rate percentiles (that is, $g_{i,t-3:t}$ in (8)), where we also allow these coefficients to differ across 5 bins formed based upon a worker's rank within the firm (defined in section 3.5).

Firms that innovate experience higher future profits, and workers in these firms experience higher earnings. By contrast, firms that fail to innovate while their competitors do experience declines in profits and worker earnings fall. Importantly, we document a significant asymmetry in our estimated profit-sharing elasticities that is consistent with the simple model we outlined in Section 1. In the model, an increase in firm innovation has a positive impact on firm profits but an ambiguous effect on worker earnings, since some workers are replaced. By contrast, an increase in competitor innovation leads to a decline in both firm profits and worker earnings.

Our findings so far do not differentiate among workers in the same firm. Hence, they correspond to the conditional mean of earnings growth faced by a particular worker employed by a given firm that, either innovates or its competitors do. In the context of our model, greater innovation by the firm also leads to higher income risk for its incumbent workers. More broadly, innovation may affect not only the conditional mean, but also the conditional variance—or higher moments—of earnings growth. For example, some innovations may introduce new production methods that displace the skills of some workers. Alternatively, firms that lose market share to competing firms may reduce their scale of production and lay off workers; if part of these workers' skills are specific to the firm, they will experience a dramatically lower growth in earnings than workers who are not laid off. Thus, focusing on average responses can mask substantial heterogeneity in ex-post outcomes across workers.

As a first step in this direction, we estimate separately equation (12) for workers that subsequently remain with the firm (stayers) or leave the firm (movers). Here, we define movers at time t as workers who have left the firm over the next 3 years; doing so allows for some delay in the decision to exit, but also allows for worker earnings at the new firm to affect the dependent variable. In the last two columns of Table 1, panel B, we re-estimate the regressions for the subset of workers depending on whether they stay or leave the firm, where we only report estimates at a 5 year horizon for brevity.

Consistent with the spirit of our model, we see that movers are more adversely affected by innovation than stayers. Earnings of workers that leave the firm do not increase following innovation by the firm—in contrast to stayers. The own firm coefficient for job stayers increases to 1.6, corresponding with an implied elasticity of 0.195, whereas the coefficient for movers is indistinguishable from zero for movers. Similarly, earnings of workers that leave the firm fall much more in response to innovation by competing firms than the earnings of workers that remain. When we reestimate competitor innovation coefficients for stayers and movers we obtain implied pass through coefficients of $0.296 \approx$ -1.5/-4.9 and $0.448 \approx$ -2.2 / -4.9 for stayers and movers, respectively. As before, however, focusing on mean effects significantly obscures heterogeneity in worker outcomes. Over the next few sections, we further elaborate on these reallocative effects.

3.4 Innovation and Worker Earnings Risk

We next examine how the *conditional distribution* of worker earnings growth rates is related to innovation by the firms, or its competitors, using quantile regressions. In particular, we next estimate the response of individual quantiles in worker growth rates $g_{i,t:t+h}$, using a specification that is directly analogous to equation (12). The only difference is that, instead of the conditional mean, we are interested in how specific percentiles of earnings growth shift in response to an innovation shock. We focus on the median, as well as six additional quantiles q describing the tails of the earnings growth distribution, $q \in \{5, 10, 25, 50, 75, 90, 95\}$. We use the functional form and methodology for jointly estimating multiple conditional quantiles of Schmidt and Zhu (2016), where the median and log of the difference between each two adjacent quantiles are assumed to follow a linear model like our specification of the conditional mean in equation (12). As before, we weigh observations by the inverse of the square root of the number of workers in each firm and compute standard errors using a block-resampling procedure that allows for persistence in the error terms at the firm level. We relegate all further methodological details to Appendix C.1.

We begin by examining how the distribution of worker earnings growth varies following innovation by the firm A_f. The top-left panel of Figure 3 plots the estimated response (the average marginal effects) of different quantiles of worker future earnings growth quantiles to firm innovation; it is the empirical analogue of Panel B of Figure 1 based on simulated data. To illustrate how the estimated marginal effects map into shifts in the distribution of earnings growth rates, the bottom-left panel compares the unconditional cumulative distribution function (CDF) of earnings growth (in black) to the implied CDF following a one-standard deviation increase in firm or competitor innovation (red line). The right side of Figure 3 plots the corresponding responses to innovation by competing firms $A_{I \setminus f}$.

In brief, we find that innovation is associated with shifts in both the variance and the skewness of future worker earnings growth. That is, the shift in average worker earnings we documented in Table 1 is distributed asymmetrically across workers. To interpret these findings, it is useful to keep in mind the rich set of control variables—which include flexible functions in worker age, and the level and growth rate in past earnings. Thus, these estimates reveal the extent to which the distribution of future earnings growth shifts for workers with similar ex-ante observable characteristics—following innovation outcomes.

Specifically, focusing on the workers employed by innovating firms, we see that a one-standard deviation increase in the firm's innovative output is associated with a 0.009 log point increase in the median earnings growth rate, which is approximately 40% smaller in magnitude than the mean responses in Table 1, suggesting substantial skewness. Indeed, we see that workers that are employed in innovating firms experience a higher likelihood of a substantial increase in their labor income: the 95^{th} and 75^{th} percentiles of income growth increase by 0.02–0.03 log points following a one-standard deviation increase in $A_{f,t}$, compared to a 0.003–0.004 log point increase in the

25^{th} and 5^{th} percentiles. Hence, the distribution of earnings growth becomes more right-skewed in innovating firms. To put these numbers in perspective, we note that the median worker in the sample experiences earnings growth of approximately zero, while the unconditional 95^{th} percentile of income growth is 0.58 log points.

The two right panels of Figure 3 examine the relation between earnings growth and innovation by other firms in the same industry. We see that workers in firms that do not innovate experience a 0.011 log point decline in their median earnings growth in response to a one-standard deviation increase in innovation by competing firms. Importantly, the distribution of earnings growth rates becomes more left-skewed as substantial earnings drops become more likely: the 10^{th} and 5^{th} percentile decrease by approximately 0.033 and 0.042 log points, respectively. These magnitudes are substantial, given that the unconditional 10^{th} and 5^{th} percentiles of cumulative earnings growth rates are -0.53 and -0.88 log points, respectively. Importantly, Appendix Figure A.4 shows these magnitudes do not substantially change across horizons of 3 to 10 years, suggesting that the effects we document represent rather permanent changes in the level of worker earnings.

In sum, we see that own-firm innovation is followed by improved outcomes for the firm's workers, whereas innovation by competing firms is associated with unambiguously worse future outcomes for these workers. To the extent that workers are risk averse, and shifts in the distribution of labor income growth represent a source of risk that they cannot diversify away, taking into account these higher-order changes in the distribution of earnings growth rates can have quantitatively different implications for the value of human capital than a pure shift in mean growth rates.

To quantify the magnitude of these estimates, we next compute the shift in the worker's utility (certainty equivalent) in response to a one-standard deviation shock in innovative output, by either the firm or its competitors. In order to do so, we need to make some assumptions. First, we assume that workers have power utility. Second, we assume that the pass-through coefficient from labor income to consumption is equal to one; lower pass-through coefficients would have the same qualitative effect as lowering risk aversion.[7] Third, to avoid extrapolating the distribution into the left tail, which will have first-order implications on worker utility, we assume that the distribution of worker earnings is characterized by 7-point distribution corresponding to the estimated percentiles. This approximation leads to more conservative estimates of utility losses, since it truncates the most extreme outcomes in the left tail (we use a more elaborate interpolation scheme in section 6 below).

We find that, even though these shifts in the distribution of future growth rates might seem modest, they have a quantitatively significant impact on worker utility. Specifically, a worker with a relative risk aversion coefficient of 5 will experience a 3.4% reduction in her utility (in certainty equivalent terms) in response to a one-standard deviation shock to innovation by competing firms—

[7] Suppose that the elasticity of an individual worker's consumption to a permanent change in income is a constant $\theta \in (0, 1)$. For an agent with a CRRA coefficient of γ, her certainty equivalent over this consumption lottery is equivalent to the certainty equivalent over the income lottery of an agent with a coefficient of risk aversion equal to $\tilde{\gamma} = 1 + (\gamma - 1)\theta$.

compared to a 1.1% reduction for a risk-neutral worker. These estimates are of the same order of magnitude than the welfare cost of business cycles due to job displacement computed by Krebs (2007). Similarly, a risk averse worker experiences a lower increase in her certainty equivalent following a one-standard deviation increase in firm innovation (0.5%), compared to a risk-neutral worker (1.3%).

3.5 Conditioning on Worker Income

A natural question is whether the effects documented above represent truly unpredictable changes in future worker earnings growth rates, or simply reflect the incomplete information of the econometrician. For instance, one possibility is that new innovations are complementary to the effort of high-skill workers, while being a substitute for low-skill workers. The increase in dispersion of future labor income growth rates we document may simply reflect an increase in the skill premium—and therefore the estimated changes in distributions may not reflect risk, but rather heterogeneous exposures across groups. Though worker skill is not observable in our data, we may expect that conditioning on workers' past earnings as a proxy for skill, might reduce the degree of dispersion in ex-post outcomes documented in Table 1.

Alternatively, we may expect the magnitude of the displacement effect to be larger for top workers, since they have the most to lose. Specifically, such is the case in the simple model in Section 1, which implies that the magnitude of the increases in the left tail should be larger for high income workers.

Here, we exploit these ideas further by estimating specifications similar to (12), except that we now allow the response coefficients to firm (a_h) and competitor (b_h) innovation to vary with the worker's current earnings rank within the firm. Following Guvenen et al. (2014), we compute worker earnings ranks based on the last 5 years of earnings—that is, using $w_{t-4,t}$, defined in equation (7).[8] Whenever we allow a_1 and a_2 to vary across groups, we also include indicator variables for each group within the specification. Also recall that, to ensure that we are not capturing the effects of mean-reversion in worker levels following a transitory shock (for instance, a bonus), we also allow the coefficients on lagged income growth rates $g_{i,t-3:t}$ to vary across firm rank bins.[9]

Figure 4 shows that conditioning on the level of worker earnings reveals greater heterogeneity in ex-post worker outcomes. Specifically, we see that the income growth rates of top-paid workers exhibit a substantial increase in dispersion (and skewness) in response to innovation than the income growth rates of lower-paid workers. Examining the top panel of Figure 4, we see that workers in the top 5 % and bottom 95% experience qualitatively similar increases in skewness in income growth

[8] As a robustness check, we also repeat our analysis by conditioning on the worker's salary rank within the industry—defined at the SIC3 level. All of our main results are similar (see Appendix Figures A.17 to A.19 for details).

[9] The astute reader will note that, given our 10% sampling rate and restriction to men only, some firm-years may not be associated with many worker observations, in which case workers' percentile ranks are not measured very precisely for small firms. To check that the potential classification errors are not driving our results, we verified that our main results hold when we drop from the estimation sample any firm-years with fewer than 20 matched workers.

rates in response to firm innovation, but the magnitudes are substantially different. For example, a one-standard deviation increase in $A_{f,t}$ is followed by a 5 percentage point increase in the 95^{th} percentile of their earnings growth rate for workers in the top 5% of the distribution, but only a 2.1 to 3.6 percentage point increase for workers in the bottom 95%.

Further, workers at or above the top 5^{th} percentile also experience a significant increase in the left tail of income growth rates following innovation by their own firm—unlike workers in the bottom 95%. This increase in the left tail dominates the location shift (increase in the median), implying that the 5^{th} and 10^{th} percentile of income growth rates actually decline for these workers. Put differently, following a higher innovative output by their own firm, highly-paid workers experience an increase in the likelihood of both large earnings gains, but also large income drops. As a result, the impact of firm innovation on worker utility is theoretically ambiguous; it depends on risk aversion. Specifically, a top-paid worker with a risk aversion of 5 experiences a 0.9% drop in her certainty equivalent following a one-standard deviation increase in A_f, compared to a 2.5% increase for a risk-neutral worker.

Highly-paid workers are also more likely to experience large income drops following higher innovation output by competing firms. Examining the bottom panel of Figure 4, we see that workers at the top 25^{th} percent experience a dramatic increase in the left-skewness of their earnings distribution compared to workers in the bottom 75^{th} percentile.[10] For instance, a one-standard deviation increase in $A_{I\setminus f,t}$ is associated with a 14 percentage point decline in the 5^{th} percentile of earnings growth for the top 5% of workers—compared to just a 1.1 2.7 percentage point fall for the workers in the bottom 75^{th} percentile. These magnitudes imply substantial utility losses following innovation by competing firms; a worker with a risk aversion coefficient of 5 experiences a 11% drop in her certainty equivalent in response to a one-standard deviation increase in $A_{I\setminus f,t}$, compared to a 2.6% drop for a risk-neutral worker.

This increased sensitivity of the earnings growth of the left tail of top workers to competitor innovation is qualitatively consistent with the simple model we outline in Section 1. In the model, top workers experience a higher proportional loss to their income when entering unemployment, which gives rise to an increased sensitivity of the left tail, especially for workers who leave the firm. In the next section, we examine the relation between innovation and worker exit in more detail.

4 Innovation and Worker Exit

Our results so far show that the distribution of worker earnings growth becomes more left-skewed following increases in the rate of innovation by competing firms. The simple model in Section 1 replicates these results through worker separations. Here, we explore the extent to which the increased likelihood of job loss helps to account for these patterns in the data.

[10] We report OLS coefficients from the same specifications in Panel A of Appendix Figure A.4.

A priori, we might expect that the likelihood of job loss accounts for most of the increase in the left tail in the data. In particular, the labor economics literature has documented that workers often experience substantial declines in earnings upon job separation (see, e.g., Jacobson, LaLonde, and Sullivan, 1993). Indeed, the same pattern holds in our data. Appendix Figure A.5 compares the distribution of earnings growth across all workers (Panel A) to those workers that remain with the same firm over the next years (Panel B) and those workers that do not (Panel C). Comparing across Panels B and C, we see that, in terms of the right tail (positive outcomes), the distribution of earnings growth looks similar across these two groups. By contrast, the left tail is substantially fatter (lower percentiles are more negative) for workers that move out of the firm relative to workers that remain with the firm. These differences are economically substantial: the magnitude of the 5^{th} and 10^{th} percentile is two to four times larger in absolute terms across all income levels for workers that move relative to those that do not.

To shed light on these issues, we next draw a distinction between workers that remain with the firm (stayers) versus those that leave (move). We then explore the extent to which the likelihood of job separation, as well as the distribution of future earnings growth conditional on job status, varies with firm innovation. Specifically, we decompose the conditional distribution of earnings growth g conditional on innovation $A = \{A_f, A_{I\setminus f}\}$ into distributions that depend on whether the worker has exited the firm ($M = 1$) or not ($M = 0$), and the probability of exit $p(M|A)$,

$$f(g|A) = f(g|A, M = 1)\, p(M = 1|A) + f(g|A, M = 0)\, (1 - p(M = 1|A))\,. \tag{13}$$

Since our data allow us to track workers across firms, we can estimate the individual components of equation (13) separately. This decomposition allows us to understand the economic drivers of the results in Figure 4—specifically, what drives the increase in the likelihood of extreme income drops in response to innovation outcomes. That is, the fact that some workers experience a higher likelihood of large income declines in response to (firm or competitor) outcomes could be purely driven by an increased likelihood of job separation. This possibility would suggest that firm-specific human capital is important. However, it is also possible that there also exists variation in worker future outcomes conditional on mobility, which would be consistent with displacement in general worker skills (e.g., a change in a worker's outside option).

4.1 Innovation and likelihood of worker exit

We begin by estimating the likelihood $p(M = 1|A)$ that a given worker remains with the same firm over the next 3 years, as a function of firm A_f and competitor $A_{I\setminus f}$ innovation. We do so by using a linear probability model,

$$M_{i,t:t+3} = a_h\, A_{f,t} + b_h\, A_{I\setminus f,t} + c_h\, Z_{i,t} + \varepsilon_{i,t}. \tag{14}$$

The left hand side variable $M_{i,t:t+3}$ takes the value 1 if employee i has a different main employer or is unemployed at $t+3$. Given our focus on 5-year income earnings changes, this choice of timing allows both for delayed mobility effects while also including the worker's earnings subsequent to exiting the firm. Using a linear model allows us to saturate the specification with the same rich set of worker- and firm-level controls Z as equation (12).

Panel A of Table 2 shows that innovative firms are more likely to retain their workers. By contrast, firms that do not innovate, while their competitors do, are more likely to lose workers at the higher end of the distribution of wages (that is, above the median). Workers at the bottom of the wage distribution are less likely to leave in response to innovation by competitors. In terms of magnitudes, a one-standard deviation increase in A_f is associated with a 1.7 to 2 percentage point lower probability that a worker leaves the firm, and the magnitudes are largely comparable across worker income levels. Conversely, a one-standard deviation increase in $A_{I\setminus f}$ is associated with a 1.8 to 4.3 percentage point higher likelihood of exit. The point estimates are larger for the workers in the top 5%, but the differences are not statistically significant. These magnitudes are substantial given that the unconditional probability that a worker leaves the firm after 3 years is 36%.

These results are largely in line with the predictions of the model in Section 1, especially when we interpret the model as applying to skilled workers. Our results are consistent with low-skill workers having fewer firm-specific skills and therefore being more easily redeployable across firms.

4.2 Innovation and earnings growth conditional on mobility

Next, we examine whether innovation is related to the distribution of future earnings growth conditional on mobility—that is, $f(g|A, M)$ in equation (13). To estimate these conditional distributions, we estimate equation (12) separately for workers that move (movers) versus those that do not (stayers). Thus, for example, $f(g|A, M = 1)$ is estimated by comparing future outcomes of workers that left the firm during periods of high innovation with outcomes for workers that left following periods of low innovation. Indeed, Figure 5 shows that the two distributions $f(g|A, M = 1)$ and $f(g|A, M = 0)$ are qualitatively different.

The first row of Figure 5 shows that higher innovation by the firm A_f is associated with an increased likelihood of substantial income declines only for exiting workers. Continuing workers experience no such increase in the left tail. The increase in the left tail among exiting workers is quantitatively larger for the highest earners, but is present across all income groups. By contrast, the response of the right tail of earnings growth—the likelihood of large earnings gains—is comparable across movers and stayers. Naturally, one interpretation of these results is that they reflect adverse selection. Specifically workers that are terminated following good shocks to the firm are more likely to be adversely selected, and therefore face worse future labor market outcomes.[11]

[11] If that is the case, we would expect to see this pattern more generally subsequent to positive firm profitability shocks. However, that does not seem to be the case. Appendix Figure A.6 shows that workers that left the firm following periods of high firm/industry stock returns do not experience more negatively-skewed income growth than

The bottom row Figure 5 shows that innovation by competing firms $A_{I\setminus f}$ is associated with an increased likelihood of large income declines for both continuing and exiting workers. However, the magnitude of the increase in the left tail is larger by a factor of two to three for exiting workers. As before, the magnitude of these effects are in general higher for more highly paid workers. Appendix Table A.6, Panel A illustrates that the same pattern is true for conditional mean (OLS) effects.

In sum, innovation is associated with variation in worker outcomes conditional on exiting the firm. By contrast, in the model in Section 1, once workers exit the firm they are all identical, that is, there is no persistent loss in skill. The fact that in the data workers that left the firm following periods of high innovation (by its competitors) fare worse relative to workers that left following periods of low innovation suggests a more persistent impact on the workers' human capital beyond the loss of firm-specific skills upon separation. We discuss this further in Section 7.

4.3 Innovation and long-term unemployment

So far, we saw that exiting workers experience a substantially more negatively-skewed distribution of earnings growth in response to innovation outcomes than continuing workers. The next step is to understand the extent to which this increased likelihood of large earnings losses are driven by longer unemployment spells or lower earnings in a new firm.

We construct a measure of long-term unemployment based on the number of years with zero W-2 earnings.[12] Specifically, $U_{i,t:t+5}$ counts the number of years between $t+1$ and $t+5$ that worker i has reported zero total earnings in her W-2 form, conditional on having left the firm she was working by time $t+3$. Table A.1 shows the distribution of $U_{i,t:t+5}$. We see that most exiting workers experience no years with zero W-2 earnings. However, there is considerable variation in the tails. Approximately 10% of exiting workers experience unemployment spells of at least a year; at least 5% of exiting workers experience unemployment spells of at least 3 years.

We estimate the following linear specification for our long-term unemployment measure,

$$U_{i,t:t+5} = a_h \, A_{f,t} + b_h \, A_{I\setminus f,t} + c_h \, Z_{i,t} + \varepsilon_{i,t}. \tag{15}$$

The vector of controls Z contains the same worker- and firm-level controls as equation (12). Panel B of Table 2 presents the results.

We find that workers that leave the firm subsequent to high innovation outcomes (either by the firm or its competitors) experience on average longer unemployment spells—as measured by years of zero reported earnings. The magnitudes are sizable: a one-standard deviation increase in $A_{f,t}$

workers that left the firm during periods of low stock returns. Likewise, Appendix Table A.6, Panel B reports OLS coefficients from the same specifications, which feature similar coefficients for competitor stock returns for both switchers and stayers.

[12] Measuring directly the length of unemployment spells is not possible in our data, since we do not observe any information on unemployment benefits. Also, note that since we exclude workers that have self-employment income in our analysis (following Guvenen et al., 2014), workers with zero W-2 earnings are not workers who switch to self employment.

is associated with 0.006 to 0.031 increase in number of years with zero W-2 earnings. Similarly, a one-standard deviation increase in competitor innovation is associated with up to a 0.038 percentage point increase in the number of years without employment; again, the magnitudes are larger for the highest-paid workers. Given that the mean number of years with zero earnings experienced by exiting workers is approximately equal to 0.33, these are economically significant magnitudes. Further, we see that the magnitudes are substantially larger for workers in the top of the earnings distribution versus the bottom.

Our results in this section shed some light on the patterns in Figure 5. Specifically, they suggest that part of the large increase in the left tail for movers following innovation outcomes we document in Figure 5 are the result of longer unemployment spells. To explore this possibility, we re-estimate the model but now excluding workers that experience any years with zero W-2 income between $t+1$ and $t+5$. Figure A.20 in the Appendix shows the results. Indeed, we find that extended periods of unemployment account for a significant fraction of the increase in the left tail following innovation outcomes. Once workers experiencing years with zero W-2 income are excluded, the increase in the left tail is still present—though significantly smaller in magnitude, especially in response to innovation by one's own firm.

A potential concern is that the measure of long-term unemployment is indirect and could simply reflect the choice to take time off work. As a more direct measure of structural unemployment, we examine worker applications for Social Security Disability Insurance (DI) benefits. One view of disability insurance is that it represents a long-term exit from the workforce, since benefits are guaranteed until medical recovery, death, or retirement at age 65 (Autor and Duggan, 2003, 2006). A potential rationale for a correlation between DI filings and our innovation measures is that some workers may satisfy medical criteria making them eligible to apply for DI, but may choose to work if opportunities in the labor market are sufficiently attractive.[13] If innovation displaces the human capital of some of these workers, we would therefore expect to see an increase in DI applications following high innovation outcomes in case labor market opportunities become scarcer for these workers.[14] Indeed, this prediction is consistent with the administrative definition of disability, which includes the 'inability to engage in a substantial gainful activity.'

To explore this further, we examine DI applications as a separate outcome. The dependent variable $D_{i,t:t+5}$ now takes the value of one if the worker i is not longer employed in the same firm she worked in at year t by the end of year $t+3$ *and* she has filed for disability insurance sometime

[13] Autor and Duggan (2003, 2006) term these workers 'conditional applicants.' Autor and Duggan discuss the secular increase in the number of 'conditional applicants' since 1984, partly as a response in changes in determination standards, but also more importantly, as changes in labor market conditions.

[14] The decision to accept transfer payments often involves exchanging a claim on a stream of (comparatively) risky future labor earnings for a known, safe stream of transfers. The cost of participation is the opportunity cost of foregone labor income, which changes in response to its riskiness. In the case of DI, some individuals who satisfy the medical criteria to claim benefit may decide whether to apply based upon changes in labor market conditions. In our model, one potential interpretation of the reservation wage b_t is the expected utility of leaving the labor force and consuming transfer income.

between year $t+1$ and $t+5$. As we see in Table A.1, applying for disability insurance is not a rare phenomenon, especially for exiting workers. Approximately 2.6% of workers apply for disability insurance over a 5-year period. Among the sample of exiting workers, this fraction rises to 4.1%.

We estimate the following linear specification,

$$D_{i,t:t+5} = a_h\, A_{f,t} + b_h\, A_{I\setminus f,t} + c_h\, Z_{i,t} + \varepsilon_{i,t}, \tag{16}$$

using the same vector of controls Z contains the same variables as equation (12).

We find that skilled workers who exit following periods of high innovation—by the firm or its competitors—are more likely to file for disability insurance relative to workers that exit following periods of low innovation (see Panel C of Table 2). These effects are present for the top of the distribution in terms of past worker earnings. Specifically, focusing on workers that are the top 75% of the distribution in terms of past age-adjusted earnings, we see that an increase in firm innovation is associated with an increase in disability applications among exiting workers. Conditional on exit, workers in the top half of the within-firm earnings distribution are also more likely to file for DI following periods of high innovation by competing firms. Interestingly, we see the opposite effect for workers at the bottom of the earnings distribution. Such workers are less likely to file for DI, conditional on exit, following high innovation outcomes by the firm. Focusing on workers in the top half of the earnings distribution, a one-standard deviation in firm (competitor) innovation is associated with a 0.06–0.09 (0.10–0.37) percentage point increase in the dependent variable.

Further, we also estimate specifications in which we do not restricting the sample to exiting workers. Estimating equations (15) and (16) for all workers, regardless of whether they left the firm or not, reveals whether innovation is unconditionally related to measures of long-term unemployment. Appendix Table A.7 shows that this is indeed the case. For example, focusing on the workers above the (firm) median earnings level, we see that a one-standard deviation increase in firm (competitor) innovation is associated with a 0.004–0.011 (0.013–0.022) percentage point increase in the number of years of zero W2 earnings (a sizable effect given a mean of 0.142). The likelihood of DI applications shows a quantitatively similar increase. We interpret these facts as additional evidence that innovation is linked with higher labor income risk, especially for the firm's top workers.

In brief, we see that most workers that exit the firm following high innovation periods have worse labor market outcomes than those who leave during low innovation periods—either by the firm or its competitors. These effects are more pronounced for workers at the top of the distribution. To some extent, these periods of unemployment help to explain the fattening of the left tail of the distribution of income growth following periods of innovation.

5 Additional Results and Robustness

Here, we discuss a number of additional results and robustness checks.

5.1 Process vs non-process

In our analysis so far, we have not differentiated among different types of innovation by the firm. In the context of the model in Section 1, innovation by the firm can have a beneficial or displacive impact on its own workers depending on whether it represents an improvement in product quality, or an improvement in production methods. Our modelling assumption is that innovation in production methods is more likely to lead to displacement for the firms' own workers. Here, we explore this idea more fully by distinguishing between process and non-process (e.g., product) innovation by the firm. Process innovation consists of new production methods or procedures that help the firm lower production costs (see, e.g., Link, 1982; Bena and Simintzi, 2019). By introducing new methods or procedures in the production process, part of the worker's human capital that was specific to the previous 'vintage' may become obsolete.

We, therefore, estimate a modified version of equation (12) that decomposes innovation by the own firm A_f into process and non-process, and examines the impact of these two types of innovation separately,

$$g_{i,t:t+h} = a_0 + a^p_\tau \, A^{proc}_{f,t} + a^o_\tau \, A^{other}_{f,t} + b_\tau \, A_{I\setminus f,t} + c \, Z_{i,t} + \varepsilon_{i,t}. \tag{17}$$

To decompose A_f into process and product innovation, we use the data and classification procedure of Bena and Simintzi (2019).[15] Bena and Simintzi (2019) identify the fraction θ_j of claims of patent j that can be identified with a process. The residual claims $1 - \theta_j$ can refer to either types of innovations, such as new products. We use these fractions to decompose the private value measure A_f into process and non-process innovations.[16] Appendix B.4 contains more details.

We expect that a process patent is more likely to be associated with displacement for the firms' own workers. Our findings are broadly consistent with this idea. The top row (Panels A and B) of Figure 6 examines how the distribution of earnings growth varies across firms that engage in process versus non-process innovation. Comparing Panels A and B, we see that the two types of innovation have a qualitatively different effect on the distribution of worker earnings growth. Product innovation is associated with earnings gains that are symmetric across workers, though higher paid workers experience a greater increase. By contrast, process innovation is associated with a substantial increase in the dispersion of earnings growth, in particular for the highest-paid workers. For these workers, a one-standard deviation increase in $A^{proc}_{f,t}$ is associated with a 4.5 percentage point decrease in the 5^{th} percentile of income growth. In Appendix Figure A.7, we demonstrate that these left tail effects for process innovation are considerably larger for workers who subsequently leave the firm, whereas estimates for non-process innovations are fairly similar between movers and

[15] Bena and Simintzi (2019) use text-based analysis to identify patent claims that refer to process innovation. In particular, they identify claims as process innovations as those which begin with "A method for" or "A process for" (or minor variations of these two strings) followed by a verb (typically in gerund form), which directs to actions that are to take place as part of the process.

[16] The correlation between A^{proc}_f and A^{other}_f is approximately 70%. A similar decomposition of $A_{I\setminus f}$ into process and non-process is somewhat less informative because the resulting series have a correlation that is in excess of 90%.

stayers.[17] Complementing our results from section 4.3, Appendix Table A.8 illustrates that, holding the total value of own firm innovation fixed, process innovation is much more predictive of long term unemployment relative to other types of innovations.

The fact that process improvements are associated with an increased likelihood of large earnings declines for the firm's top workers might be surprising, if one has the view that process improvements mostly displace low-skilled (and hence low-earning) workers. However, this need not be the case; process innovations are often associated with significant organizational changes, and often lead to the replacement of mid-level executives that lack the skills, or willingness, to adapt to new production methods.[18] That said, not all process innovations may be truly novel. In Appendix Figures A.23 and A.24, we thus consider one alternative text-based measure of the extent to which the innovation is novel to the firm. Specifically, we use the backward-similarity measure of Kelly, Papanikolaou, Seru, and Taddy (2020) to further separate patents into two categories based on whether the text is similar ('less novel') or fairly distinct ('novel') from prior patents received by the firm. Analogous to changes in process innovation, we find that more novel innovations (especially novel process innovations) by the firm are more likely to be associated with fattening of the left tail for high income workers.

5.2 Robustness

The results in Figure A.2 suggest that our innovation measures do not pick up underlying, pre-existing trends in firm-specific profitability. However, innovation could be related to a time-varying unobservable firm characteristic that also drives profits. For instance, firms that hire 'good' executives may decide to invest more in innovation and at the same time grow faster; such firms may also have a different wage structure than firms with 'bad' managers. In the absence of a randomized experiment that allows us to assign different innovation outcomes across different firms, we have performed the following alternative estimation strategies. Specifically, we have expanded the set

[17] Appendix Figure A.11 repeats this exercise allowing for coefficients to vary with worker tenure as well as mobility. We find that these left tail effects for high income workers are most pronounced for those with 3 or more years of tenure with the firm, especially those who leave.

[18] For instance, one of the best-selling management texts (Davenport, 1993, p.194) on process innovation advocates: *[...]Interventions are needed to maximise gains and prevent backsliding [such as] replacing resisters and/or individuals who have failed to adapt to the new environment. Leaders of successful process change replace resisters and those who cannot adapt to the new environment only after providing education, training, and coaching, and allowing them ample time to adapt.* On discussing the implementation of process innovation in the Distributed Systems Manufacturing (DCM) Group, which was a part of Digital Equipment Corporation (DEC): *In 1985, the DSM team developed an aggressive 5-year plan. A systems and information management-tools component called for the implementation of computer-aided design, computer-integrated manufacturing, artificial intelligence, group technologies and other advanced manufacturing systems, many of which had significant impacts on how people in the organization worked.* (Davenport, 1993, pp.168-170). Indeed, our data support this view. During the 1986–1988 period, DEC engaged in significant process innovation (according to our A_f^{proc} measure, that put it at the top 10–25% for all firms engaging in (non-zero) process innovation in each year. Further, these process improvements were also associated with substantial turnover of mid-level managers: *DSM's group manager, like most successful process change leaders, used a combination of hard and soft interventions to manage anticipated resistance. [...]But the group manager also displayed the impatience for results that is characteristic of successful change leaders, and did not hesitate to replace resisters and others whom he felt were not adapting quickly enough.* (Davenport, 1993, p.195).

of covariates that we include in the vector Z to include firm-level decision variables related to innovation—specifically, controls for the ratio of R&D spending to book assets. When doing so, we are essentially comparing firms that spend the same money on R&D, but have different innovation outcomes. Figure A.15 in the Appendix shows that we find that controlling for R&D spending does not really affect the magnitude of the estimated coefficients.

The results in Figure 4 suggest that, highly paid workers likely face substantially higher income risk in response to innovation than lower-paid workers. However, a potential caveat with this interpretation is that these results may be driven by differences in how employees exercise stock options. Specifically, in contrast to other forms of compensation, the gains from stock option grants appear in the worker's W-2 form when these options are exercised by the employee, rather than when they are granted to the worker. One possibility is that, for a given innovation outcome, the firm always grants top employees the same amount of stock options. However, if these employees exercise these options at different points in time, their capital gains will vary—and we may therefore see a greater dispersion in their ex-post income growth rates. Though this may indeed be a possibility, our results suggest that it is unlikely to be a key driver of our findings. First, the increase in the left tail seems to be driven by workers who leave the firm, rather than those who stay and likely receive more option compensation. Second, the results are consistent when we focus on outcomes that do not rely on option exercise gains—specifically, the number of years of zero earnings or the likelihood of applying for DI (Table 2). Third, as we discuss next, the fattening of the left tail that we document in Figures 4 appears to be specific to innovation outcomes.

If the increase in the dispersion of earnings growth rates is due to differences in the timing that options are exercised, we should see a similar pattern in response to shocks to firm profitability more generally. We perform two sets of comparisons. First, we re-estimate equation (12), but now replace the firm and competitor innovation measures (A_f and $A_{I\setminus f}$) with the firm's own stock return R_f, and the value-weighted stock return of the other firms in the same industry $R_{I\setminus f}$ in that year. Using stock returns instead of changes in realized profits has the advantage that it includes changes in firm future profitability. Indeed, Vuolteenaho (2002) documents that most of the variation in firm-level returns can be attributed to shocks to current and future profitability ('cashflow news'). Panels A and B Figure 7 presents the results. Second, we repeat this exercise where we instead condition on realized profit growth over the next five years, both by the firm (Panel C) and its competitors (Panel D). In the case of the firm, our measure is identical to the dependent variable used in Figure A.2, and we aggregate competitor innovation analogously.

Contrasting Figure 7 to Figure 4, we see that the relation of the earnings of top workers with innovation outcomes is qualitatively distinct than the relation with other sources of firm profitability.[19] In Panel A, we see that a positive shock to the firm's stock price is associated

[19]We also present OLS coefficients in Appendix Table A.4 and compare them against comparable estimates for innovation. Consistent with the argument in Kline et al. (2019), pass through coefficients for own firm innovation are higher relative to those associated with stock returns or changes in profits. Appendix Table A.6 reports analogous

with mostly a symmetric increase in worker earnings across the wage distribution. In Panel B, the difference is even starker. Specifically, controlling for the firm's own stock return, an increase in the stock market valuation of competing firms is associated with a weakly *positive* effect on worker earnings growth—in contrast to the increased likelihood of sharp income declines we saw in Figure 4. This comparison suggests that differences in the timing of employee stock options are unlikely to be responsible for the increase in dispersion—and particularly, the left tail—of earnings growth for top workers.[20] Panels C and D paint a similar picture; notably, in panel D, we observe that competitor profit growth is associated with fairly modest improvements in worker outcomes in the left tail for low skill workers. Point estimates for the right tail of the distribution are small and generally statistically insignificant. More generally, our interpretation of this comparison is that the creative destruction aspect of innovation is an important source of differentiation from other shocks to firm profitability (e.g., shocks to product demand) in how firm shocks impact employee earnings.

Last, another potential concern is that the higher left-tail sensitivity for higher income workers is really picking up an age effect. That is, even though worker earnings are defined net of lifecycle dummies—see equation (7)—this only adjusts for mean effects. Workers at the right tail of the distribution in terms of wages could be older. Older workers can have higher exposure to innovation if they are less able to adapt to new production methods. To explore this idea further, we estimate equation (12) separately for workers of different age groups and income levels. Indeed, Appendix Figure A.13 shows that younger workers fare somewhat better than older workers in response to innovation outcomes by either the firm or its competitors. However, our main finding that top workers are more exposed to innovation, especially in terms of the left tail of income growth, continues to apply within age categories.[21]

6 Implications for Income Inequality

Several studies have documented a significant increase in earnings inequality over the last two decades (Piketty and Saez, 2003). Importantly, much of the recent increase in earnings inequality is a between-firm phenomenon (Song et al., 2019; Barth et al., 2016). To the extent that worker earnings are related to firm profits and innovation is related to firm growth, a natural candidate explanation for these patterns is changes in innovation patterns across firms. Indeed, we saw in Figure 2 that innovation outcomes became much more disparate in the 1990s relative to the 1980s; this is primarily a within-industry phenomenon that was partially reversed in the 2000s. Here,

estimates conditional on mobility for innovation and stock returns, respectively.

[20] Appendix Figure A.6 further confirms our conclusion: there is no analogous fattening of the left tail for workers who move following periods with high firm stock returns.

[21] Likewise, we also allow for heterogeneous effects by prior earnings and tenure at the firm, where were split workers into two categories based on whether they have 3 or more years of tenure (the median in the full sample) or 5 or more years of tenure (the median in the matched sample of workers at public firms). Our results, which are presented in Appendix Figures A.9 to A.12, continue to hold within tenure groups, and displacement effects are somewhat more pronounced for higher tenure workers (especially those that move).

we explore how these shifts in the distribution of firm innovation translate into changes in the distribution of worker earnings levels. In general, the direction of the effect is ambiguous: we may expect that an increase in the dispersion of innovation outcomes across firms leads to greater inequality across workers; however, also recall that top earners display somewhat greater sensitivity of earnings to firm or competitor innovation outcomes. The fact that top workers exhibit greater sensitivity to competitor innovation that workers and the bottom of the earnings distribution could lead to a drop in inequality in response to greater dispersion in innovation outcomes across firms—if it turns out to be the case that the displacement effect of top workers is particularly strong.

We provide a quantitative answer to this question using the point estimates from our quantile regressions. We briefly summarize the main steps of the procedure here; we relegate all details to Appendix C.2. We begin with an individual's earnings level measured over the prior 3 years—net of age effects, specifically $w^i_{t-2,t}$, the base rate for our growth rate construction in equation (8). We then simulate a counterfactual earnings growth rate for each worker—conditional on income levels using the point estimates from Section 3.5—under the assumption that all firms in the same industry always innovate by the same amount. That is, we eliminate within-industry dispersion in firm innovation outcomes at time t. Given these simulated growth rates and the initial level of earnings $w^i_{t-2,t}$, we compute average worker earnings levels over the next 5 years $w^i_{t+1,t+5}$ for each worker. Last, we compute the level of inequality in average worker earnings over the next five years $w^i_{t+1,t+5}$ under these counter-factual realizations for A_f and $A_{I\setminus f}$ at time t. By comparing the actual levels of income inequality with these counterfactual levels, we can assess how disparities in firm innovation contribute to worker earnings inequality at each point in time.

Figure 8 presents our findings. We examine separately inequality in the top (the 95/50 range) and the bottom (the 50/5) range of the distribution in the left and right panels, respectively. We plot both the realized change in income inequality (orange line) and the change that our estimates attribute to changes in the distribution of innovation outcomes documented in Figure 2. Focusing on Panel A, we see that our point estimates imply that the increase in dispersion in firm innovation outcomes led to an acceleration of the growth in inequality at the top during the 1990–2000 period. By contrast, the model assigns only a minor role to between-firm dispersion in generating changes in inequality at the bottom of the income distribution. Examining Panels B and C, we see that the increased dispersion in innovative outcomes significantly contributed to increased between-firm income inequality at both the top and the bottom of the distribution. By contrast, it mostly contributed to higher within-firm inequality at the top of the distribution (the 95/50). As we see in the right figure of Panel C, within-firm inequality at the bottom is mostly driven by the economic cycle.

In terms of magnitudes, our calculations imply that dispersion in firm innovation outcomes accounts for most of the increase in inequality at the top during this period. In particular, our simulations imply that if all firms in the same industry were innovating by the same amount, the increase in overall income inequality at the top would be 90% smaller than its realization; by contrast, its effect on inequality at the bottom of the income distribution is much smaller (12%). In terms of between-firm inequality, our simulations imply that eliminating firms dispersion in innovation outcomes would have led to a 60% to 70% smaller increase in inequality at the top and at the bottom of the distribution. Last, our results indicate that the increased dispersion in firm innovation outcomes can account for 80% of the increase in within-firm inequality at the top, while only 20% at the bottom.

In sum, we find that the asymmetry in innovation outcomes across firms in the 1990s potentially contributed significantly to the observed rise in income inequality. That said, one potential source of concern in interpreting these findings stems from the fact that $A_{f,t}$ is somewhat persistent over time—its serial correlation is approximately 0.7. In this case, we may be worried that we are overestimating the impact of a shock in a given year: our quantile regressions are not merely picking up the effect of $A_{f,t}$ on $w^i_{t+1,t+5}$, but also the effect of $A_{f,t+1}, A_{f,t+2}\dots$ etc. In this case, summing up over the estimated effect of $A_{f,t}$ over time would tend to overestimate their overall impact on inequality. We think this is a valid concern, but unlikely to significantly impact our findings: focusing on average earnings over a period $w^i_{t+1,t+5}$ overweighs the impact of $A_{f,t}$ relative to $A_{f,t+1}$ etc. Further, the fact that the point estimates of the response of $w^i_{t+1,t+k}$ on $A_{f,t+1}$ are essentially identical for $k = 3$ and $k = 5$ (see Figure A.4) suggests that our quantile regressions are not excessively estimating the impact of $A_{f,t}$ on average worker earnings.

7 Discussion

The simple model we outlined in Section 1, captures many of the key features in the data. However, it has little to say about the persistence of earnings losses as a result of innovation. The model is essentially static in that workers' displacement is summarized by (temporary) unemployment. There is no impact of innovation on a workers' ability to find another job. By contrast, several aspects of our empirical analysis suggest a role for technology permanently displacing workers' existing skills. First, consider the variation in outcomes among exiting workers. Figure 5 and Table 2 show that workers that leave the firm following high innovation outcomes (by either the firm or its competitors) experience worse outcomes in terms of the distribution of earnings growth and duration of unemployment relative to workers that leave the firm during low innovation periods. Second, worker earnings often respond to increases in innovation by competing firms in the industry, even if these shifts are unrelated to firm profits. Specifically, Appendix Figures A.21 and A.22 show that impactful (highly-cited) patenting activity by other firms is also related to worker earnings—even though the link with firm profits is significantly weaker (Kogan et al., 2017). To account for these effects the model would have to be modified to allow for vintage-specific human capital, as in the models Chari and Hopenhayn (1991), Jovanovic (1998) and Violante (2002). Such a model is outside the scope of this paper.

Can our results be explained with alternative models that do not feature skill displacement? A model of adverse selection (Gibbons and Katz, 1991) may potentially help interpret some—but not all—of the facts we document. Specifically, the fact that competitor innovation has an adverse effect on worker outcomes (Figure 4) may simply reflect the fact that the firm reduces overall employment as a result (Table 1); if the firm knows more about the worker's skill than the labor market, being fired is a negative signal about worker ability, and would therefore lead to perhaps longer unemployment spells or a reduction in her wage in a new firm. However, for adverse selection to explain the contrast between Figure 4 and Figure 7, it has to be the case that innovation is somehow different than other shocks which affect the firm. One possibility is that when a firm innovates, it somehow reveals information about its current workers (e.g., about some new dimension of skill that was previously unused) that leads to permanent differences in earnings ex-post. If other firms in the same industry use the same technology, terminating a worker will signal to the other firms that he lacks skills that have become relevant.

8 Conclusion

Our analysis reveals several new facts regarding how the distribution of worker earnings growth changes following technological improvements by firms (and its competitors). In general, we find that innovation by the firm is associated with higher worker earnings, though the gains are asymmetrically distributed. By contrast, innovation by competing firms is associated with mostly lower worker earnings, and, more importantly, an increase in the likelihood of large income declines. We draw two broad conclusions from our findings.

First, workers at the top of the earnings distribution display substantially higher sensitivity to either firm or competitor innovation outcomes, suggesting that these workers bear significant labor income risk, a point emphasized by Parker and Vissing-Jorgensen (2009) and Guvenen et al. (2014), among others. The fact that innovation is typically associated with an increased possibility of substantial earnings loss for these workers represents a source of risk that cannot be easily diversified. Workers may require higher wages, or alter their portfolio composition in response. Second, our simulations that are consistent with innovation playing a significant role in driving the increase in income inequality during the 1990s. Importantly, the main effect appears to have been not the increase in the level of aggregate innovative activity during this period, but rather the fact that this increased amount of innovation was concentrated on a small subset of firms.

By linking detailed information about technological advances of firms and their competitors with administrative earnings data, our analysis helps to deepen our understanding of the fundamental drivers of individual-level income dynamics. These estimates inform policy and welfare analysis for several reasons. First, changes in earnings risk immediately directly affect projections of the fiscal impact of a variety of government programs. For instance, Social Security and Medicare revenues and future payments depend nonlinearly on earnings; so a fiscal assessment requires a full estimate of the distribution of earnings and its evolution over time. Second, changes in expected future earnings levels and risk can change incentives for individuals to claim program benefits. To this end, we find that the same groups of workers whose earnings risk appears to increase following new technological developments are also more likely to apply for Social Security Disability Insurance benefits, which is consistent with the "conditional applicants" hypothesis of Autor and Duggan (2003). Finally, optimal design of social insurance programs, including public pension, unemployment insurance, and disability benefits, critically depends on the quantity and fundamental sources of uninsurable income risk.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Acemoglu, D. and P. Restrepo. Robots and jobs: Evidence from us labor markets. *Journal of Political Economy forthcoming*.

Aghion, P., U. Akcigit, A. Bergeaud, R. Blundell, and D. Hemous (2018). Innovation and Top Income Inequality. *The Review of Economic Studies 86*(1), 1–45.

Aghion, P., A. Bergeaud, R. Blundell, and R. Griffith (2017). Innovation, firms and wage inequality. Working paper, Harvard University.

Aghion, P. and P. Howitt (1992, March). A Model of Growth through Creative Destruction. *Econometrica 60*(2), 323–51.

Autor, D., D. Dorn, G. H. Hanson, and J. Song (2014). Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics 129*(4), 1799–1860.

Autor, D., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Autor, D. H. and D. Dorn (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review 103*(5).

Autor, D. H. and M. G. Duggan (2003). The rise in the disability rolls and the decline in unemployment. *The Quarterly Journal of Economics 118*(1), 157–206.

Autor, D. H. and M. G. Duggan (2006). The growth in the social security disability rolls: A fiscal crisis unfolding. *Journal of Economic Perspectives 20*(3).

Autor, D. H., L. F. Katz, and M. S. Kearney (2006, May). The Polarization of the U.S. Labor Market. *American Economic Review 96*(2), 189–194.

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the united states. *Journal of Labor Economics 34*(S2), S67–S97.

Bena, J. and E. Simintzi (2019). Machines could not compete with chinese labor: Evidence from US firms' innovation. *Working Paper*.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly Journal of Economics 128*(3), 967–1015.

Chari, V. V. and H. Hopenhayn (1991). Vintage human capital, growth, and the diffusion of new technology. *Journal of Political Economy 99*(6), 1142–1165.

Constantinides, G. M. and A. Ghosh (2017). Asset pricing with countercyclical household consumption risk. *The Journal of Finance 72*(1), 415–460.

Davenport, T. (1993). *Process Innovation: Reengineering Work Through Information Technology.* Harvard Business Review Press.

Evans, D. S. (1987, August). Tests of Alternative Theories of Ind Growth. *Journal of Political Economy 95*(4), 657–74.

Fagereng, A., L. Guiso, and L. Pistaferri (2018). Portfolio choices, firm shocks, and uninsurable wage risk. *The Review of Economic Studies 85*(1), 437–474.

Garleanu, N., L. Kogan, and S. Panageas (2012). Displacement risk and asset returns. *Journal of Financial Economics 105*, 491–510.

Gibbons, R. and L. F. Katz (1991). Layoffs and lemons. *Journal of Labor Economics 9*(4), 351–380.

Goldin, C. and L. Katz (2008). *The Race Between Education and Technology.* Harvard University Press.

Goldin, C. and L. F. Katz (1998). The origins of technology-skill complementarity. *The Quarterly Journal of Economics 113*(3), 693–732.

Goos, M. and A. Manning (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The Review of Economics and Statistics 89*(1), 118–133.

Guiso, L., L. Pistaferri, and F. Schivardi (2005). Insurance within the firm. *Journal of Political Economy 113*(5), 1054–1087.

Guvenen, F., S. Ozkan, and J. Song (2014). The Nature of Countercyclical Income Risk. *Journal of Political Economy 122*(3), 621–660.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics 36*(1), pp. 16–38.

Hornstein, A., P. Krusell, and G. L. Violante (2005). The Effects of Technical Change on Labor Market Inequalities. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth.* Elsevier.

Hornstein, A., P. Krusell, and G. L. Violante (2007). Technology—Policy Interaction in Frictional Labour-Markets. *Review of Economic Studies 74*(4), 1089–1124.

Howell, S. T. and J. D. Brown (2020). Do cash windfalls affect wages? evidence from r&d grants to small firms. Working paper, NYU.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993). Earnings losses of displaced workers. *The American Economic Review 83*(4), 685–709.

Jovanovic, B. (1998). Vintage Capital and Inequality. *Review of Economic Dynamics 1*(2), 497–530.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2020). Measuring technological innovation over the long run. *American Economic Review: Insights forthcoming.*

Kline, P., N. Petkova, H. Williams, and O. Zidar (2019). Who profits from patents? rent-sharing at innovative firms. *The Quarterly Journal of Economics 134*(3), 1343–1404.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Kogan, L., D. Papanikolaou, and N. Stoffman (2020). Left behind: Creative destruction, inequality, and the stock market. *Journal of Political Economy 128*(3).

Krebs, T. (2007). Job displacement risk and the cost of business cycles. *American Economic Review 97*(3).

Lagakos, D. and G. L. Ordoñez (2011). Which workers get insurance within the firm? *Journal of Monetary Economics 58*(6), 632–645.

Lamadon, T., M. Mogstad, and B. Setzler (2019). Imperfect competition, compensating differentials and rent sharing in the us labor market. *Working Paper*.

Link, A. N. (1982). A disaggregated analysis of industrial r&d: Product versus process innovation. *The Transfer and Utilization of Technical Knowledge. D. Sahal (ed.)(Lexington, MA: DC Heath)*.

Papanikolaou, D. (2011). Investment shocks and asset prices. *Journal of Political Economy, forthcoming*.

Parker, J. A. and A. Vissing-Jorgensen (2009). Who bears aggregate fluctuations and how? *American Economic Review 99*(2).

Piketty, T. and E. Saez (2003). Income inequality in the united states, 1913–1998. *The Quarterly Journal of Economics 118*(1), 1–41.

Schmidt, L. D. and Y. Zhu (2016). Quantile spacings: A simple method for the joint estimation of multiple quantiles without crossing.

Schmidt, L. D. W. (2016). Climbing and falling off the ladder: Asset pricing implications of labor market event risk. Working paper, University of Chicago.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *The Quarterly Journal of Economics 134*(1), 1–50.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature 49*(2), 326–65.

van Reenen, J. (1996). The creation and capture of rents: Wages and innovation in a panel of u. k. companies. *The Quarterly Journal of Economics 111*(1), 195–226.

Violante, G. L. (2002). Technological acceleration, skill transferability, and the rise in residual inequality. *The Quarterly Journal of Economics 117*(1), 297.

Vuolteenaho, T. (2002). What drives firm-level stock returns? *The Journal of Finance 57*(1), 233–264.

Tables and Figures

Table 1: Innovation, firm profitability and worker earnings

	A. Firm Profitability			B. Worker Earnings				
Horizon (years)	(3)	(5)	(10)	(3)	(5)	(10)	(5)	
				All Workers			Stayers	Movers
Firm Innovation, market value (A_f^{sm})	6.81	7.99	8.82	1.38	1.38	1.07	1.56	0.03
	(8.71)	(7.39)	(6.17)	(15.46)	(11.33)	(10.01)	(11.80)	(0.28)
Implied Elasticity				0.203	0.173	0.121	0.195	0.004
Competitor Innovation, market value ($A_{I\setminus f}^{sm}$)	-3.94	-4.93	-5.99	-1.45	-1.88	-2.28	-1.46	-2.21
	(-7.85)	(-7.81)	(-5.19)	(-5.42)	(-8.45)	(-9.27)	(-5.85)	(-7.91)
Implied Elasticity				0.368	0.381	0.381	0.296	0.448
R^2	0.197	0.220	0.233	0.045	0.050	0.054	0.122	0.079

Note: Table reports the relation between firm innovation, firm profitability and worker earnings, specifically point estimates of equation (11) and (12) in the main text. The table relates firm profitability and worker earnings to innovation by the firm (A_f, defined in equation (9) and the innovation by the firm's competitors ($A_{I\setminus f}$, the average innovation of other firms in the same SIC3 industry, see equation (10)). For the firm-level regressions (Panel A), controls include one lag of the dependent variable, log values of firm capital, employment, and the firm's idiosyncratic volatility, and industry (I) and time (T) fixed effects. All firm-level variables are winsorized at the 1% level using annual breakpoints. Standard errors are clustered by firm and year. All right-hand side variables are scaled to unit standard deviation. The worker-level regressions (Panel B) use the specification described in equation (12), using the same weighting and block-subsampling inference procedure as the quantile regression specifications. The last two columns estimate (12) at the worker level using the subsample of stayers and movers, where movers are defined as workers who leave the firm within the next three years. Please see the main text for further details.

Table 2: Innovation, mobility, and long-term unemployment

A. Indicator for leaving the firm within 3 yrs ($\times$100)

Innovation	Worker wage rank				
	[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	-1.91	-1.74	-1.74	-1.70	-1.46
	(-14.65)	(-13.63)	(-13.13)	(-12.80)	(-9.56)
Innovation by competitors, $A_{I\setminus f}$	-1.09	-0.20	0.42	1.11	1.13
	(-4.22)	(-0.79)	(1.67)	(4.33)	(4.03)

B. Number of years unemployed ($\times$100), 5yr horizon
(conditional on having left firm within 3 yrs)

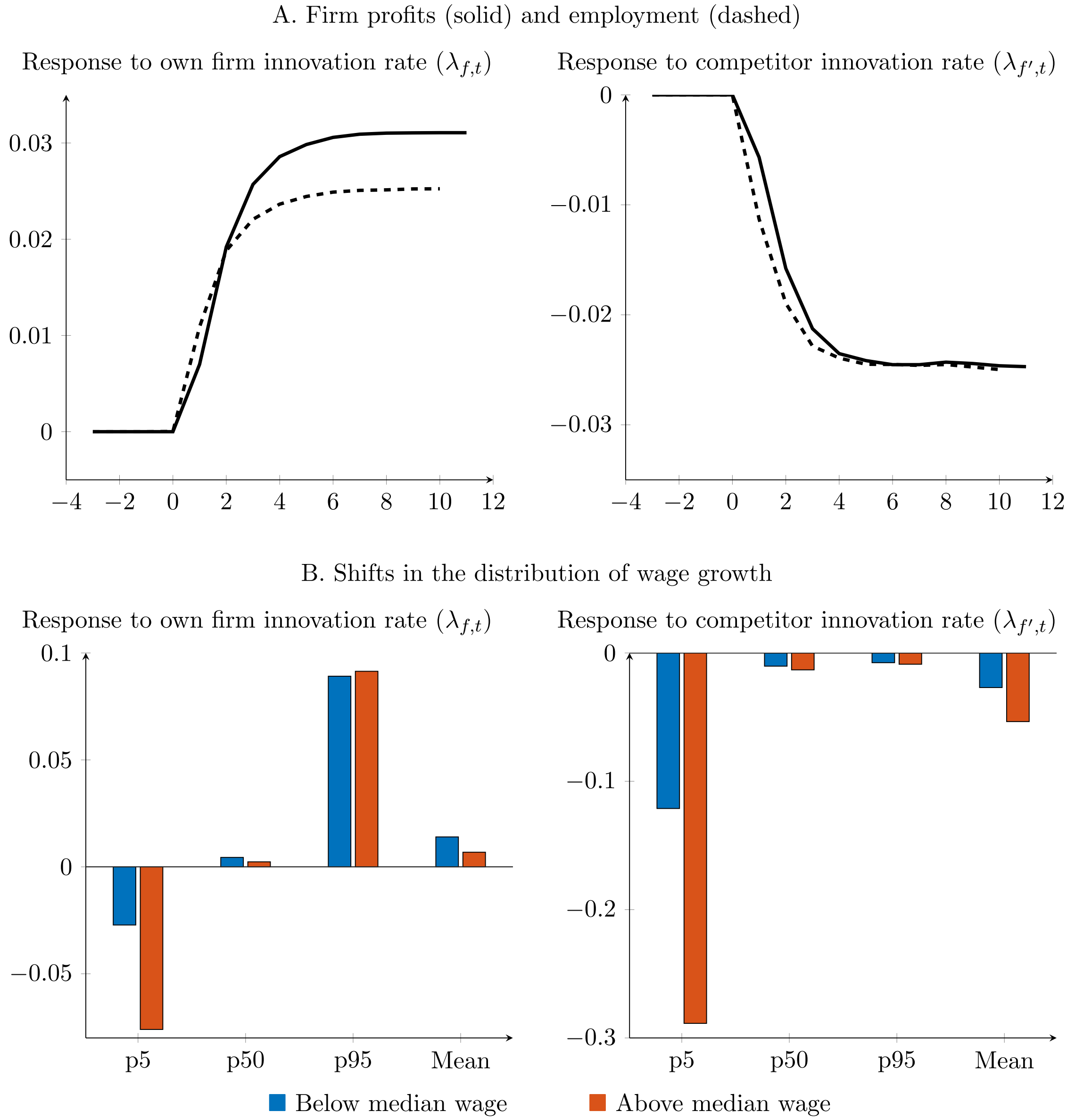
Innovation	Worker wage rank				
	[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	0.62	1.72	1.48	1.86	3.07
	(2.93)	(8.46)	(7.58)	(6.31)	(6.72)
Innovation by competitors, $A_{I\setminus f}$	-0.44	1.24	2.92	3.87	3.83
	(-1.28)	(4.33)	(10.20)	(10.78)	(8.04)

C. Application for disability insurance (DI) indicator ($\times$100), 5yr horizon
(conditional on having left firm within 3 yrs)

Innovation	Worker earnings rank				
	[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	-0.05	0.06	0.06	0.07	0.09
	(-1.73)	(2.51)	(3.72)	(5.22)	(5.02)
Innovation by competitors, $A_{I\setminus f}$	-0.37	-0.16	0.10	0.32	0.37
	(-6.14)	(-3.40)	(2.54)	(8.15)	(6.87)

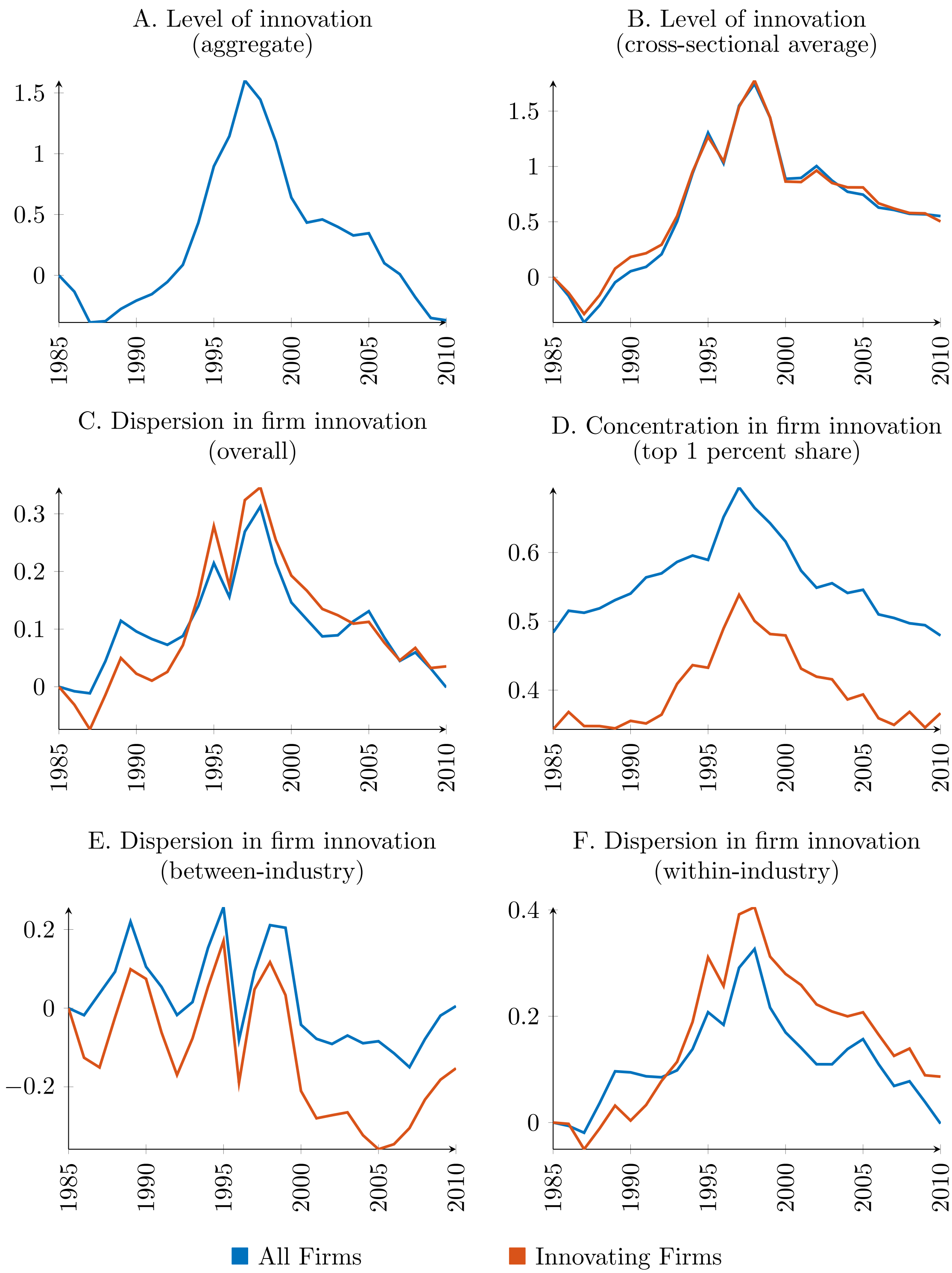
Note: Panel A reports point estimates of OLS regressions of equation (14) in the paper. The dependent variable is a dummy ($\times$100) which equals 1 if, at $t+3$, a worker is no longer employed at the same firm as at time t. Panel B reports OLS estimates of equation (15) in the paper. The dependent variable is a count of the number of years of zero W2 earnings worker i has experienced between years $t+1$ and $t+5$ ($\times$100), conditional on having left the firm by year $t+3$. Panel C reports estimates of equation (16) in the paper. The dependent variable is a dummy that takes the value of 1 if worker i has applied for disability insurance at any point between years $t+1$ and $t+5$ ($\times$100), conditional on having left the firm by year $t+3$. In all cases, we allow the response of the dependent variable to innovation (by the firm A_f or its competitors $A_{I\setminus f}$) to vary based on the worker's earnings rank, which are defined net of deterministic life-cycle effects. The coefficients are standardized to a unit-standard deviation shock in the independent variable. Standard errors, in parentheses, are clustered at the firm level.

Figure 1: Model: Response to an increase in the rate of innovation



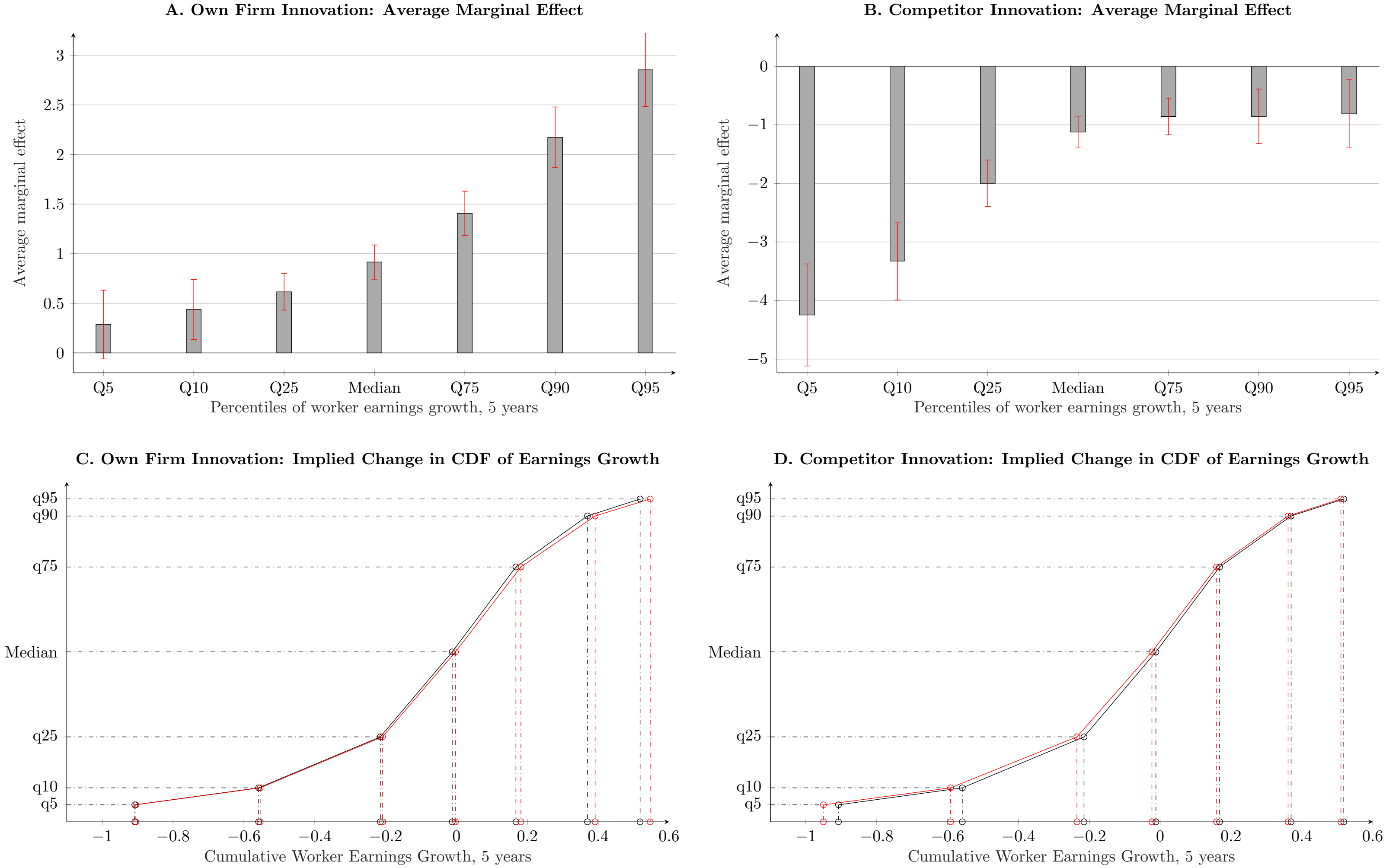
Note: The left figure of Panel A plots the response of firm profits (solid line) and employment (dashed line) to a positive shock to the rate of firm innovation, $\lambda_{f,t}$—a switch to $\lambda_{f,t} = \lambda_H$ at time 0. The figure on the right plot the corresponding response to a shock to the innovation rate of competing firms $\lambda_{f',t}$ at time 0. Panel B plots the response of the distribution of future earnings growth over the next year for workers that are currently employed at firm f. The left figure plots the response to the firm's own innovation ($\lambda_{f,t}$) while the right plots the response to innovation by its competitor($\lambda_{f',t}$). For each model simulation, we simulate a path for firm profits, employment and worker earnings for $t = -100$ to $t = 10$. We simulate two firms f and f' and set the maximum number of product lines (to be shared across these two firms) to 100. We set $\mu = 0.6$; $\lambda_L = 0.2$; $\psi = 0.5$; $\bar{h} = 1$; $\lambda_H = 2$; $\chi = 0.25$; and $p = 0.8$. We set unemployment earnings b_t to be equal to 75% of the minimum quality level across goods in the economy. The model is simulated in continuous time ($dt = 1/200$) and then time-aggregated to annual observations. Analogous to panel A, we compare various summary statistics – the 5th, 50th, and 90th percentiles, as well as the mean – of the distribution of wage changes experienced by individual workers between a situation where $\lambda_{f,t}$ (left panel) or $\lambda_{f',t}$ switches from λ_L to λ_H at $t + dt$ versus a counterfactual in which the rate of innovation does not change. We do this separately for workers with below- (blue columns) and above-median (orange columns) wages at the focal firm, respectively. The outcome variable is the model analogue of the main outcome variable used in our empirical analysis: the logarithmic growth rate of the next 5 years of workers' earnings relative to the past 3 years of earnings.

Figure 2: Firm Innovation during our sample period



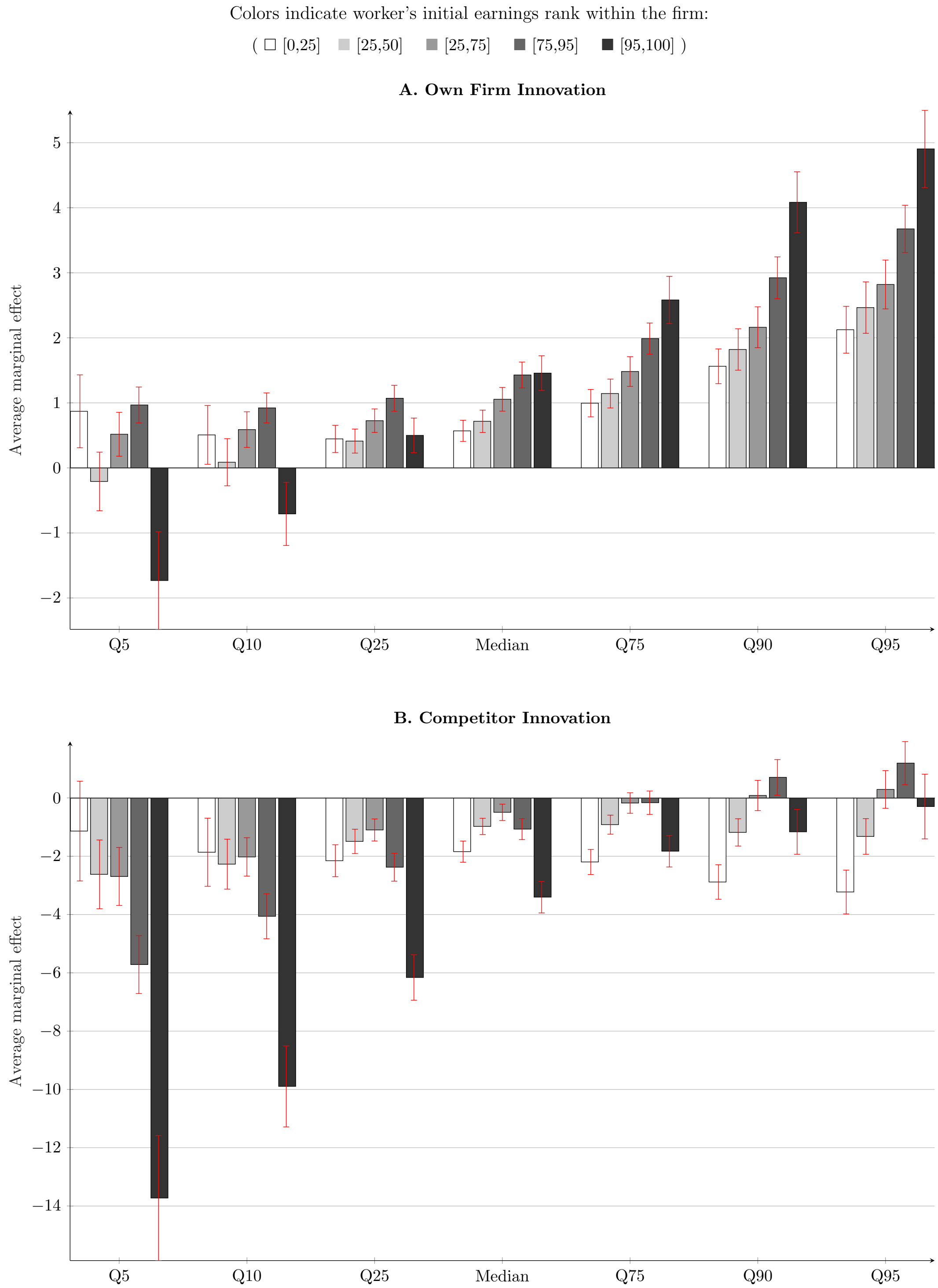
Note: Figure plots the distribution of firm innovation over time. We report results both for the entire sample as well as for firm-year observations with non-zero values of innovation in a given year (that is, among innovative firms). Panel A plots the aggregate level of innovation, defined as in Kogan et al. (2017): $\sum_f \sum_{j\in P_{f,t}} \xi_j / \sum_f B_{ft}$. In Panel B, we report the cross-sectional mean of A_f. Panel C plots the dispersion in firm innovative activity, defined as the coefficient of variation in A_f (that is, the ratio of the cross-sectional standard deviation (overall, between- and within-industry) scaled by the mean of innovation A_f in each year). Panel D plots the share of innovation (in terms of patent market values ξ_j) that is accounted for by the top 1% of firms in each sample. Panels E and F report the dispersion in the between- and within-industry component of A_f, respectively. Industries are defined at the SIC3 level. All series are in log deviations from their 1985 levels.

Figure 3: Earnings growth and firm innovation



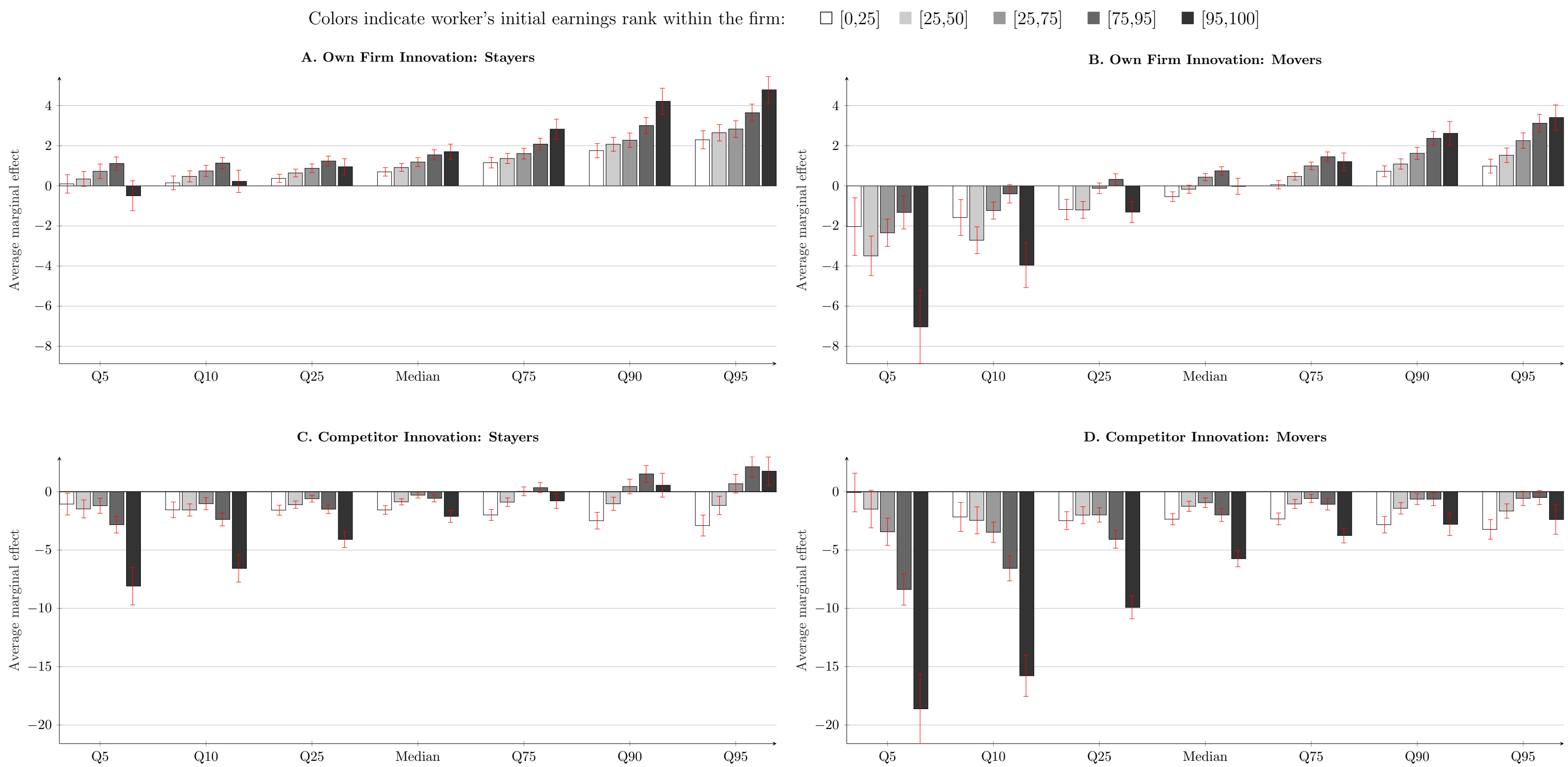
Note: The top panel plots the average marginal effects of 5-year worker earnings growth in response to a one-standard deviation increase in firm and competitor innovation that are implied by the quantile regression estimates (equation (12) in the main text) across different horizons. The units on the vertical axis correspond to log points (times 100). The bottom panel plots the implied change in the cumulative distribution of worker earnings growth in response to firm and competitor innovation: the black line denotes the unconditional distribution, and the red line denotes the distribution subsequent to a one-standard deviation shock.

Figure 4: Earnings growth and innovation conditional on earnings levels



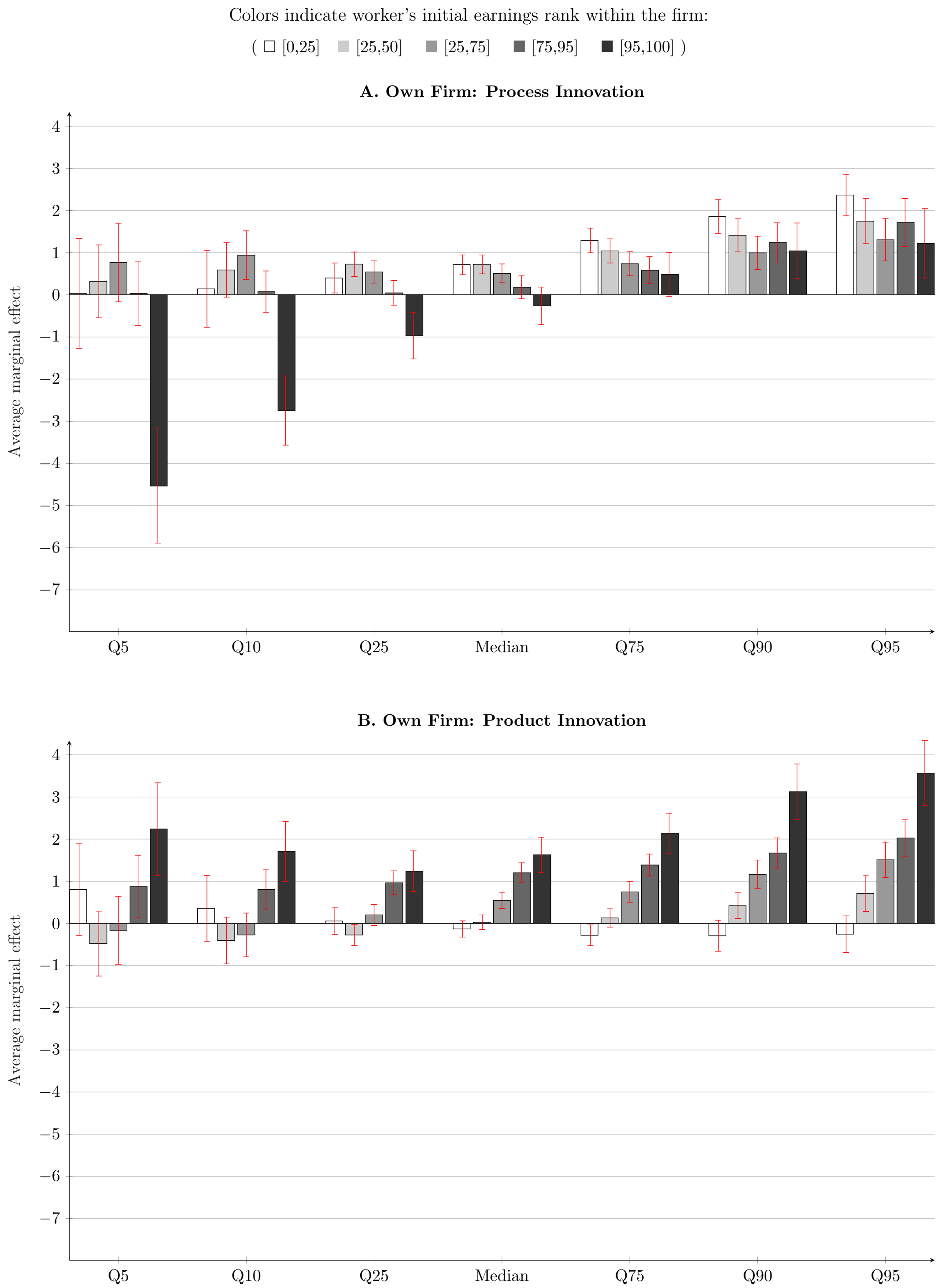
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text), where we allow for heterogenous coefficients for workers with different earnings levels and estimates are scaled to correspond with a 1 standard deviation change in each variable of interest. The worker earnings rank is defined net of deterministic life-cycle effects. We focus on 5-year growth rates. The units on the vertical axis correspond to log points (times 100).

Figure 5: Earnings growth and innovation: movers versus continuing workers



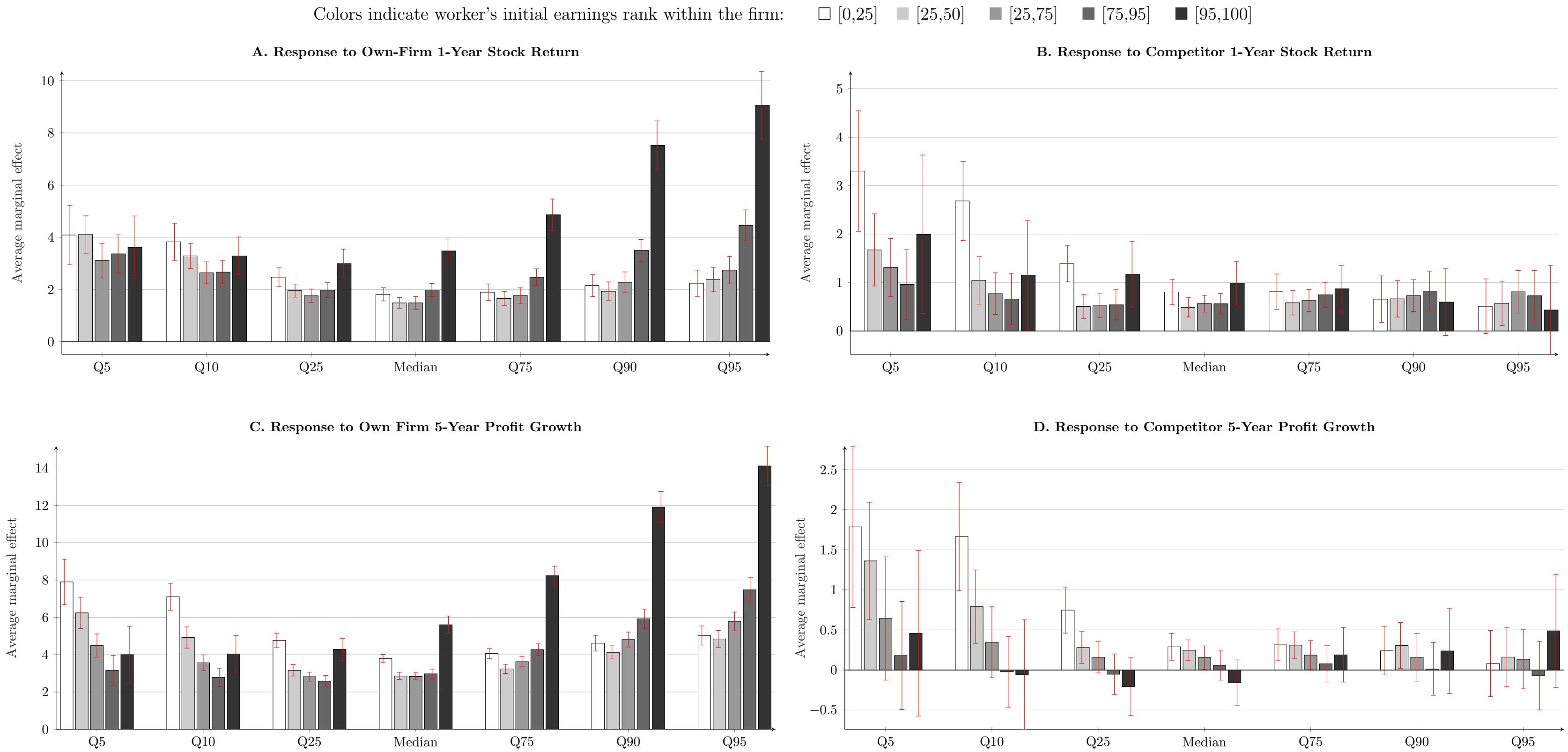
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels, where estimates are scaled to correspond with a 1 standard deviation change in each variable of interest. The equation is estimated separately for workers that remain with the firm (stayers) versus workers that leave the firm (switchers). The worker earnings rank is defined net of deterministic life-cycle effects. We focus on 5-year growth rates. The units on the vertical axis correspond to log points (times 100).

Figure 6: Earnings growth and own firm innovation, process vs non-process



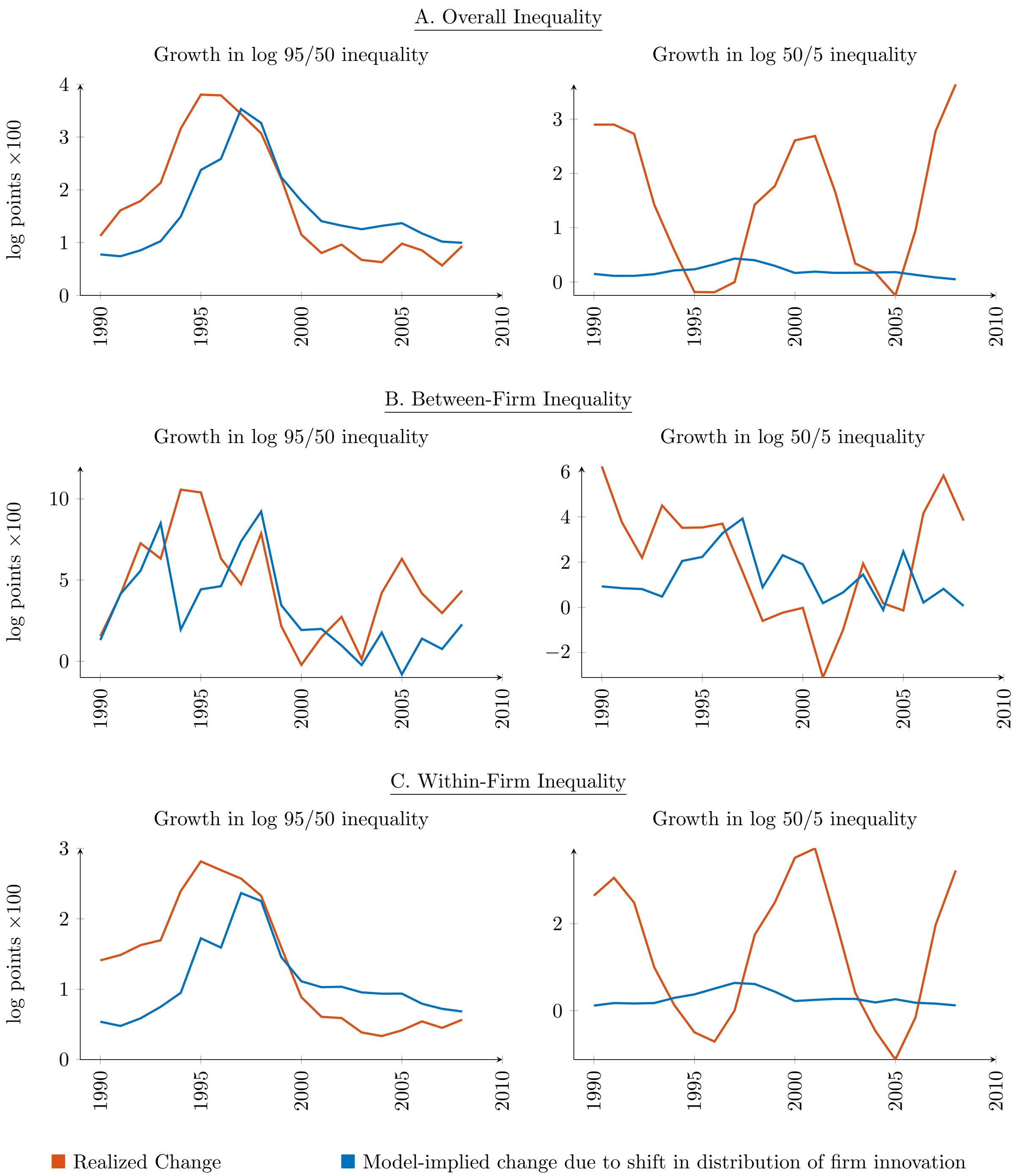
Note: Fgure plots the average marginal effects own firm process and non-process oriented innovations that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. The worker earnings rank is defined net of deterministic life-cycle effects. We focus on 5-year growth rates. Coefficients are scaled to correspond with a 1 unit standard deviation change in each of the independent variables. The units on the vertical axis correspond to log points (times 100).

Figure 7: Earnings growth and other firm/competitor outcomes: stock returns and profit growth



Note: The top panel of the figure plots the average marginal effects of firm—and competitor—stock returns that are implied by the quantile regression estimates (analogous to equation (12) in the main text, except that we use own firm and competitor year t returns in place of the innovation measures) for workers with different earnings levels, where own firm and competitor innovation measures are replaced by own firm and competitor stock returns in the same year. Estimates are scaled to correspond with a 1 standard deviation change in each variable of interest. Bottom panel reports average marginal effects from an alternative specification where own firm and competitor innovation are replaced with realized own firm and competitor cumulative profit growth over the next 5 years. The worker earnings rank is defined net of deterministic life-cycle effects. We focus on 5-year growth rates. The units on the vertical axis correspond to log points (times 100).

Figure 8: Changes in the Distribution of Worker Earnings



Note: The figure relates recent trends in wage inequality to firm innovation through simulations of our quantile regression model. We focus on two measures of earnings inequality—the 95–50 and the 50–5 spread in log earnings—that capture the right and left tail. The orange line at time t corresponds to actual increase in inequality from time t to $t+5$ (divided by 5). The blue line uses the point estimates from the quantile regressions in Section 3.5 to estimate the shifts in inequality in worker earnings that can be attributed to changes in inequality in innovation outcomes across firms (see, e.g. Figure 2). Specifically, we first compute the counterfactual level of worker earnings inequality that would obtain in a world in which all firms in the same industry innovated equally at a given point in time (we set $A_{f,t}$ to equal its industry-year mean) using the actual worker-level residuals. The blue line at time t corresponds to the change between time t and $t+5$ in the difference between realized inequality and the counterfactual, which corresponds to the role of inequality in firm innovation outcomes at time t in generating inequality during that period. For additional details, see Appendix C.2.

A Analytical Derivations

Here, we briefly describe the derivation of the model. The demand for variety i can be obtained from the optimization problem of the final-goods producer

$$\max_{\{x_{i,t}\}} \left(\int_0^1 x_{i,t}^{\nu} \right) - \int_0^1 p_{i,t}\, x_{i,t}, \tag{A.1}$$

which yields the inverse demand curve as the first-order condition for good i,

$$\begin{aligned} p_{i,t} &= \nu\, x_{i,t}^{\nu-1}. \\ \left(\frac{p_{i,t}}{\nu}\right)^{\frac{1}{\nu-1}} &= x_{i,t} \end{aligned} \tag{A.2}$$

Next, consider the problem faced by the leading producer of good i. She can set a price such that the firm with the next-best level of efficiency–denoted by q'–finds it unprofitable to produce. That is, she sets a limit price equal to the second-best producer's marginal cost,

$$p_{i,t} = \frac{W_t}{q'_{i,t}}, \tag{A.3}$$

where W is the marginal cost of employing the fixed factor l. Hence, the leading firm produces an amount equal to

$$\left(\frac{W_t}{\nu q'_{i,t}}\right)^{\frac{1}{\nu-1}} = x_{i,t}. \tag{A.4}$$

Hence, the flow profits to the leading producer are equal to

$$\begin{aligned} \Pi_{i,t} &= \left(p_{i,t} - \frac{W_t}{q_{i,t}}\right) x_{i,t} \\ &= (\kappa - 1) \left(\frac{\kappa}{\nu}\right)^{\frac{1}{\nu-1}} \left(\frac{q_{i,t}}{W_t}\right)^{\frac{\nu}{1-\nu}}, \end{aligned} \tag{A.5}$$

where $\kappa \equiv q_{i,t}/q'_{i,t} > 1$ is the efficiency gap. These profits include the payment to the skilled worker/manager. Since she can only steal a fraction β if she diverts one unit of output, the solution to this contracting problem is to give her a fraction β of the profits, in which case she is indifferent between stealing versus not. Last, market clearing determines W_t

$$\begin{aligned} 1 &= \int_0^1 l_{i,t}\, di \\ W_t &= \frac{\nu}{\kappa} \left[\int_0^1 q_{i,t}^{\frac{\nu}{1-\nu}}\, di \right]^{1-\nu}. \end{aligned} \tag{A.6}$$

B Data Appendix

Here, we describe our data construction in more detail.

B.1 SSA administrative earnings records

For our empirical analysis, we work with a 10% random sample of confidential, panel earnings records for males which is drawn from the U.S. Social Security Administration (SSA)'s Master Earnings File (MEF). The MEF includes annual earnings information which is top-coded at the SSA annual contribution limit prior to 1978, and uncapped information on annual earnings from 1978-2013. Due to several potential measurement issues in the initial years following the transition to uncapped earnings, we start our analysis in 1980 (see, e.g., Guvenen et al., 2014, for further details).

Our main sample selection criteria and variable construction methods, unless otherwise stated, closely follow Guvenen et al. (2014). Specifically, we exclude the self-employed and exclude worker-years for individuals who have not had at least three out of the prior five years of earnings exceeding a minimum threshold. The minimum earnings threshold is the amount one would earn working 20 hours per week for 13 weeks at the federal minimum wage. For an individual worker to appear in the sample at time t, we require that she not receive self-employment earnings in excess of 10% of total wage income or the above minimum earnings threshold in any of the years which are used to construct either conditioning or dependent variables. All earnings are converted to 2010 dollars using the personal consumption expenditure deflator. We restrict attention to workers who are above the age of 25 at time t, and, when we calculate growth rates, we require that the worker has at least one year with earnings above the threshold during which he is below the age of 60. Even after applying these filters, the sample includes over 100 million worker-year observations.

In addition to total annual earnings, the MEF also includes detail on the Employer Identification Number (EIN) and SIC codes of the three employers which were associated with the highest annual earnings for each individual. This information allows us to link each worker-year earnings measure with a particular firm and industry, and also to detect when workers switch employers. When an individual receives income from more than one job in a given year, we associate her with the EIN of the firm that pays the highest total wage, following Autor et al. (2014) and Song et al. (2019).

Using this mapping between workers and EINs, we also construct measures of an individual's tenure within the firm, which is the number of consecutive years for which the firm has been the worker's largest source of W-2 income. In addition, we construct measures of switches between firms. For instance, we can compute the probability that a worker who is currently employed at firm j at time t continues to be employed at the firm at time $t+h$.

Following Guvenen et al. (2014), we estimate age dummy coefficients by regressing log wages on age and cohort-specific effects in a random 10% sample of the data. We choose 25 year old as the omitted base category, so the age dummy captures the average ratio of the log wage of an older worker to a 25 year old over the sample period. $D(\text{age}_{i,t})$ is obtained by exponentiating the age dummies. At the very start of a worker's income record, we only divide by dummy variables associated with years after his first W-2 record above the minimum earnings threshold. For example, if a worker who is 25 at time t had his first W-2 record above the minimum threshold in time $t-2$, then we only divide through by $\sum_{j=0}^{2} D(age_{i,t-j})$.

In addition to total annual earnings, the MEF also includes detail on the Employer Identification Number (EIN) of the three employers which were associated with the highest annual earnings for each individual. This information allows us to link each worker-year earnings measure with a particular firm and industry, and also to detect when workers switch employers. Using this mapping between workers and EINs, we also construct measures of an individual's tenure within the firm, which is the number of consecutive years for which the firm has been the worker's largest source of W-2 income.

B.2 Constructing a Matched Sample of Public Firms

Next, we use the EIN numbers in the MEF to map the innovation measures–which are only available for public firms–with individual workers' earnings histories and to get a richer picture of how the conditional distribution of workers' income growth rates change with innovative output. For instance, the EIN appears directly below the legal company name on the cover page of the annual (10-K) and quarterly (10-Q) financial statements. Following standard practice, we exclude from the analysis financial firms (SIC codes 6000-6799) and utilities (SIC codes 4900-4949), as well as firms for which cannot find EINs, which leaves us with a sample of around 142 thousand firm-years over the 1980-2013 period. While SIC code information is available in the MEF, we use the SIC code information from their financial statements for our analysis.

We combine two data sources in order to match the firm identifier (GVKEY) from CRSP-Compustat Merged (CCM) database to EINs to the MEF. First and foremost, EIN numbers of publicly traded firms are readily available in their SEC filings, appearing on the front page of each firm's annual report (form 10-K). We can access both current and historical EIN information from the company header files, which gives us a set of EINs which are associated with a given firm. In a small number of cases, the same EIN can be associated with multiple firm identifiers (GVKEYs). In the vast majority of cases, only one of the two records is active over a given date range, or one of the two filers is a subsidiary of the other. In the latter case, we associate the EIN with the GVKEY of the parent firm. In the small remainder of cases, we only keep the GVKEY-EIN mapping from the current header file. However, the EIN from the 10-K may only be picking up a subset of the total employee base for each of these firms, because many firms pay workers through multiple EINs. For instance, Song et al. (2019) report that, according to Dun & Bradstreet data, the average firm listed on the New York Stock Exchange is associated with 3.2 EINs. To this end, the gap between the employment measure from firms' 10-K (which also includes employment in other countries and subsidiaries) and the number of W-2's in the MEF tends to be largest in percentage terms for the firms with the highest reported 10-K employment.

To improve our coverage of employment at firms with multiple EINs, we bring in an additional source of information. We augment our existing list of GVKEY-EIN links with information from firms' form 5500 filings, which are publicly-available documents that report information about firms' benefit plans to comply with the Employee Retirement Income Security Act (ERISA).[22] This dataset provides a link between company identifying information (name, address, etc.) and EINs, and includes approximately 600 thousand unique EIN numbers per year starting in 1999. Prior to 1999, filings by firm plans with fewer than 100 participants are not included in the FOIA data, so sample sizes are smaller. We can then link company names on form 5500 to a list of "major subsidiaries" in Exhibit 21 which each firm is required to file on its annual report. Combining these two sources allows us to associate a given GVKEY with additional EINs of firm subsidiaries and/or other EINs associated with parent firms' retirement plans. We are extremely grateful to Josh Rauh and Irina Stefanescu for sharing a link file between the form 5500 data and CCM data which was used in Rauh and Stefanescu (2009) and Rauh, Stefanescu, and Zeldes (2019), which we used as a starting point for the empirical analysis.

Incorporating subsidiary information increases the size of our estimation sample by about 50%, from 7.8 million to 11.4 million worker-years in our baseline estimation. However, we do note that our main results appear unchanged if restrict the sample to using the EIN numbers from the current header file only, which corresponds with the EIN a firm's most recent 10-K.

[22] We access FOIA information for filings from 1999-present from the US Department of Labor's website: https://www.dol.gov/agencies/ebsa/about-ebsa/our-activities/public-disclosure/foia/form-5500-datasets. Information from 1990-1998 is taken from the Center for Retirement Research at Boston College University: http://crr.bc.edu/data/form-5500-annual-reports/.

Figure A.1, panel A, shows the number of public firms with EINs which are matched and unmatched to W-2 records in the MEF by year. On average, matching rates are quite high. We can find records in the MEF for about 84% of the public firm-years. That said, there is a core group of around 650 firms that we cannot find per year, which causes overall matching rates decline to some extent post-2000 due to a gradual decline in the total number of public firms. Figure A.1, panel B, shows the number of matched and unmatched firms by major SIC industry group. We observe that the industry composition of the two samples are broadly similar.

Table A.1 provides summary statistics for observations which meet our screening criteria for being included in the estimation for the full sample and matched sample, respectively. We note that employees at public firms are slightly older and earn about $16 thousand dollars more per year. Workers at matched public firms have about a year of additional tenure on average, and are also more likely to have tenure greater than or equal to 3 years relative to workers at non-matched firms. Recall from our earlier discussion that the tenure measure is censored by the fact that our sample starts in 1980; therefore, these summary statistics provide a lower bound on the population distribution of firm tenure. For this reason, our empirical specifications involving tenure will emphasize a binary measure which is not subject to this downward bias.

Table A.3 compares the characteristics of matched and unmatched firms over our sample period. Matched firms tend to be similar in terms of book assets, but larger in terms of employment (as reported on 10-K forms). Given the discussion above about the fact that some firms may have multiple EINs associated with different divisions and/or subsidiaries, such a result is to be expected. Matched firms are also somewhat more innovative. The ratio of R&D to assets, as well as average values of each of our three innovation measures–which we will describe in the next section–are all higher for the sample of matched firms. For the sample of matched firms, we can also compute a measure of total employment from the SSA data by counting up the total number of W-2's associated with each employer. The average firm in our matched sample has about 3,800 employees according to this measure. On average, the SSA-implied employment measure is smaller than the number reported in firms' financial statements. This result is unsurprising given that the 10-K number is more inclusive and the fact that some firms may pay employees through multiple EINs, not all of which are found in the 5500 data.

B.3 Cumulative Earnings Growth and Transitory Shocks

When focusing on income changes, our main variable of interest will be the growth in age-adjusted income $w^i_{t,t+k}$ over a horizon of h years, defined as follows:

$$Y_{i,t:t+h} \equiv w^i_{t,t+h} - w^i_{t-2,t}. \tag{A.7}$$

Here, we have chosen as our baseline the average (age-adjusted) earnings between $t-2$ and t as the scaling factor; that said, our results are similar if we extent the window to 5 years. Focusing on the growth of average income over multiple horizons in (A.7) has two distinct advantages. First, summing over multiple years yields a much smaller number of observations with zero income relative to a simple comparison of year-on-year income changes. Second, and more importantly, this transformation can smooth out some large changes in earnings that may be induced by large transitory shocks, which places a higher emphasis on persistent earnings changes. See Appendix B.3 for more details.

To see the second point more clearly, suppose that annual log income, net of age effects, is the sum of a random walk component $\xi_{i,t} = \xi_{i,t-1} + \eta_{i,t}$ plus an i.i.d transitory component $\varepsilon_{i,t}$. In our benchmark specifications, we set $h = 5$; in this case, a log-linear approximation of our five year earnings measure $Y_{i,t:t+h}$ around zero is:

$$\begin{aligned} Y_{i,t:t+5} &\approx \frac{1}{5}\eta_{i,t+5} + \frac{2}{5}\eta_{i,t+4} + \frac{3}{5}\eta_{i,t+3} + \frac{4}{5}\eta_{i,t+2} + \eta_{i,t+1} + \frac{2}{3}\eta_{i,t} + \frac{1}{3}\eta_{i,t-1} \\ &+ \frac{1}{5}\left[\varepsilon_{i,t+5} + \varepsilon_{i,t+4} + \varepsilon_{i,t+3} + \varepsilon_{i,t+2} + \varepsilon_{i,t+1}\right] - \frac{1}{3}[\varepsilon_{i,t} + \varepsilon_{i,t-1} + \varepsilon_{i,t-2}]. \end{aligned} \tag{A.8}$$

That is, our transformation implicitly computes a weighted average over permanent and transitory shocks of different periods. Our measure places a larger weight on the short-term permanent shocks (e.g., on η_{t+1}) than in the long-term shocks (e.g., at η_{t+5}). More importantly, however, the transitory shocks ε receive mostly a lower weight than the permanent shocks η, hence reducing their importance. Last, given that our measure essentially is an equal-weighted average over the transitory shocks, it is likely to be closer to a normal distribution if the underlying ε shocks are non-normally distributed.

B.4 Measuring Innovation

Here, we summarize the main steps behind the construction of the innovation measure, and refer the reader to Kogan et al. (2017) for additional details.

The Kogan et al. (2017) estimate of the economic value of patent j equals the estimate of the stock return due to the value of the patent times the market capitalization M of the firm that is issued patent j on the day prior to the announcement of the patent issuance:

$$\xi_j = (1 - \bar{\pi})^{-1} \frac{1}{N_j} E[v_j | r_j] \; M_j. \tag{A.9}$$

An important step in this construction is the estimation of the conditional expectation $E[v_j|r_j]$. Kogan et al. (2017) allow for the possibility that the stock price of innovating firms may fluctuate during the announcement window for reasons unrelated to innovation, and hence include an adjustment for measurement error that requires parametric assumptions. We follow their methodology closely. Next, part of the value of the patent may already be incorporated into the stock price, hence (A.9) includes an adjustment that is a function of the unconditional probability $\bar{\pi}$ of a successful patent application—which is approximately 56% in the 1991-2001 period (see, e.g., Carley, Hegde, and Marco, 2014). Since this adjustment does not vary by patent, it has no impact on our analysis. Last, if multiple patents N_j are issued to the same firm on the same day as patent j, we assign each patent a fraction $1/N_j$ of the total value.

The next step involves aggregating (A.9) at the firm and industry level. To construct the measure at the firm level, we sum up all the values of patents $j \in P_{f,t}$ that were granted to firm f in calendar year t,

$$\xi^{sm}_{f,t} = \sum_{j \in P_{f,t}} \xi_j. \tag{A.10}$$

In addition to the measures of innovation based on stock market reactions (A.10), we also construct a measure that weigh patents by their forward citations. Specifically, we measure the amount of innovation by firm f in year t as

$$\xi^{cw}_{f,t} = \sum_{j \in P_{f,t}} \frac{1 + C_j}{1 + \bar{C}_j}, \tag{A.11}$$

where $\bar{C}_j$ is the average number of forward citations received by the patents that belong in the same technology class (as measured by 3-digit CPC codes) and were granted in the same year as patent j. This scaling is used to adjust for citation truncation lags (Hall et al. (2005)) as well as differences in citation patents across technology classes. Both (A.10) and (A.11) are essentially weighted patent counts; if firm f files no patents in year t, both variables are equal zero.

Large firms tend to file more patents. As a result, both measures of innovation above are strongly increasing in firm size (Kogan et al., 2017). To ensure that fluctuations in size are not driving the variation in innovative output, we scale the measure above by firm size. We use book assets as our baseline case,

$$A^k_{f,t} = \frac{\xi^k_{f,t}}{B_{ft}}, \qquad k \in \{sm, cw\}. \tag{A.12}$$

We note that our main results are not sensitive to using book assets for normalization since we also control for various measures of firm size in all our specifications. Our main results are similar if we scale by the firm's market capitalization instead.

We also construct a measure of innovation by competing firms. We define the set of competing firms as all firms in the same industry – defined at the SIC3 level– excluding firm f. We denote this set by $I \setminus f$. We then measure innovation by competitors of firm f as the weighted average of the innovative output of its competitors,

$$A^k_{I\setminus f,t} = \frac{\sum_{f' \in I \setminus f} \xi^k_{f',t}}{\sum_{f' \in I \setminus f} B_{f't}}, \qquad k \in \{sm, cw\}. \tag{A.13}$$

To decompose the total value of innovation into these two types, we rely on the data and classification procedure of Bena and Simintzi (2019). Bena and Simintzi (2019) use text-based analysis to identify patent claims that refer to process innovation. In particular, they identify claims as process innovations as those which begin with "A method for" or "A process for" (or minor variations of these two strings) followed by a verb (typically in gerund form), which directs to actions that are to take place as part of the process. Hence, once can identify the fraction θ_j of claims of patent j that can be identified with a process. The residual claims $1 - \theta_j$ can refer to either types of innovations, for example, new products. We use these fractions to decompose the private value measure A_f into process and non-process innovations; Appendix B.4 contains more details.

To construct $A^{proc}_{f,t}$ and $A^{other}_{f,t}$, we use a similar procedure as equations (A.10) and (A.12). We create an estimate of the dollar amount of process innovation by the firm in year t as

$$\xi^{proc}_{f,t} = \sum_{j \in P_{f,t}} \theta_j \, \xi_j, \tag{A.14}$$

as well as the residual innovation $\xi^{other}_{f,t} = \xi_{f,t} - \xi^{proc}_{f,t}$. Similar to equation (A.12), we scale both measures by firm assets. In terms of magnitudes, the average fraction of the dollar value of firm innovation that can be characterized as process, $\xi^{proc}_{f,t}/\xi_{f,t}$, is approximately 27.5%.

C Methodology

Here, we relegate details of our estimation and simulation methodology.

C.1 Econometric Methodology

Quantile regression methods are semiparametric, allowing us to characterize features of conditional distributions without needing to fully specify distributional assumptions. Just as OLS regression methods estimate best linear projections of conditional expectation functions under misspecification, linear quantile regression methods estimate a (weighted) linear approximation of the true unknown quantile function. See Angrist, Chernozhukov, and Fernández-Val (2006) for further details. In contrast to alternative parametric methods for characterizing higher moments of non-Gaussian distributions (e.g., fitting mixture models), quantile regression methods are highly computationally tractable. We estimate the parameters of interest by solving a sequence of convex optimization problems which converge quickly even with a large number of observations (14.6 million) and conditioning variables (our baseline specification includes hundreds of regressors).

In our analysis, we use a method for estimating multiple conditional quantiles recently developed in Schmidt and Zhu (2016). This method, which is a natural extension to the location-scale paradigm, has the advantage of estimating conditional quantiles which are not susceptible to the well-known quantile crossing problem. Furthermore, as we will show in the next section, it allows for a natural interaction between aggregate and cross-sectional determinants of higher moments. In what follows, we briefly describe the procedure, and refer the reader to Schmidt and Zhu (2016) for more details.

Let $Y_{i,t}$ be the dependent variable of interest, and $X_{i,t}$ be a set of observable conditioning variables. In our case, $Y_{i,t}$ will be the growth rate of labor income, cumulated over various horizons. Let $q(\alpha; x)$ be the conditional quantile function of $Y_{i,t}$, for each $\alpha \in (0, 1)$, satisfying

$$q(\alpha; x) \equiv \inf\{y \in \mathbb{R} \colon P[Y_{i,t} \leq y \ \mid \ X_{i,t} = x] \geq \alpha\}. \tag{A.15}$$

If we further assume that the distribution of $Y_{i,t}$ is absolutely continuous, then $q(\alpha; x)$ is a continuous, strictly increasing function of α. Our interest will be in estimating a model for p conditional quantiles associated with the probability indices $\alpha_1, \ldots, \alpha_p$, and we will denote the j^{th} conditional quantile of interest by $q_j(x) = q(\alpha_j; x)$. We assume throughout that $\alpha_{j^*} = \frac{1}{2}$ for $j^* \in \{1, \ldots, p\}$, so $q_{j^*}(x)$ is the conditional median of $Y|X = x$.

Following Schmidt and Zhu (2016), we parameterize the conditional quantiles $q_j(x)$ by:

$$q_j(x) = \begin{cases} x'\beta_0 & \text{if } j = j^* \\ x'\beta_0 - \sum_{k=j}^{j^*-1} \exp(x'\beta_k) & \text{if } j < j^* \\ x'\beta_0 + \sum_{k=j^*+1}^{j} \exp(x'\beta_{k-1}) & \text{if } j > j^* \end{cases}. \tag{A.16}$$

The econometric model in (A.16) is a natural extension of the location-scale paradigm. All quantiles are anchored to the conditional median of $Y|X$ — which is denoted by $q_{j^*}(x)$. The quantiles above the median are estimated by adding nonnegative functions ("quantile spacings") which are exponentially affine in the independent variables $X_{i,t}$, which ensures that all quantiles will be properly ordered (e.g., the 75-th percentile will always be above the median).[23]

Our specification for multiple quantiles allows for considerable flexibility in higher moments above and beyond the location-scale benchmark in the previous section. For instance, spacings to the left of the median could be larger than spacings to the right of the median, which would indicate a left-skewed distribution of shocks. Or, alternatively, some variables could have a larger influence on more extreme spacings (such as the distance between the 10-th and 5-th percentile) relative to spacings closer to the median (such as the distance between the median and the 25-th percentile), generating variation in conditional kurtosis. Moreover, Schmidt and Zhu (2016) argue that a multiplicatively separable functional form like (A.16) can be motivated by the nonparametric extension of differences-in-differences estimation proposed by Athey and Imbens (2006).

The interpretation of an individual slope coefficient within one of the spacing functions is a

[23] Any sequence of conditional quantiles of an absolutely continuous random variable can be decomposed as a median plus or minus a sequence of non-negative distances of quantiles. For computational tractability, we require that the specification is *linear in parameters*. Schmidt and Zhu (2016) demonstrate how to estimate the model in (A.16) by iteratively applying a sequence of standard linear-in-parameters quantile regressions, beginning with the median and working toward the tails.

semi-elasticity. In particular, for any $j \neq j^*$, we have that

$$\beta_j = \frac{\partial}{\partial x}[\log(q_{j+1}(x) - q_j(x))], \tag{A.17}$$

which is the percentage change in the distance between two quantiles induced by a marginal change in x. A positive slope coefficient in a spacing below the median ($j < j^*$) indicates that, all else constant, increasing x increases downside risk, fattening the left tail. Positive coefficients in spacings above the median are associated with a fattening of the right tail.

We present our main results in terms of the average marginal effect of the independent variables x on a given quantile. These estimates incorporate the accumulated effect across quantiles. An advantage of our estimation methodology is that it results in highly tractable forms for these average marginal effects,

$$E\left[\frac{\partial q_j(X_{i,t})}{\partial X_{i,t}}\right] = \begin{cases} \beta_0 & \text{if } j = j^* \\ \beta_0 - \sum_{k=j}^{j^*-1} E[\exp(X'_{i,t}\beta_k)]\beta_k & \text{if } j < j^* \\ \beta_0 + \sum_{k=j^*+1}^{j} E[\exp(X'_{i,t}\beta_{k-1})]\beta_k & \text{if } j > j^* \end{cases}. \tag{A.18}$$

To estimate these average marginal effects, we use the sample means as plug-in estimators of the expectations. In some of our specifications, the particular coefficient of interest is an interaction term of a categorical variable with some other continuous variable (e.g., innovation). In these cases, we compute an average marginal effect for the subsample of workers within that category.

We compute standard errors using a block-resampling procedure that allows for persistence in the error terms at the firm level. Schmidt and Zhu (2016) establish the consistency, asymptotic normality, and consistency of a bootstrap inference procedure. For computational efficiency, we use a subsampling procedure rather than the bootstrap, noting that subsampling methods are generally valid under weaker conditions than the bootstrap. We estimate the variance-covariance matrix of the unknown vector of parameters by randomly selecting 10% firms without replacement, then scaling the variance-covariance matrix of the subsampled parameters appropriately using the asymptotic rate of convergence of the ($\sqrt{N}$-consistent) estimator. We also stratify these firm subsamples by 10 size bins. We use 100 replications.

To circumvent the incidental parameter bias, we exclude very small industries from the analysis. Specifically, we drop industries with less than 10,000 matched worker-year observations from the estimation. We impose the same restriction in the OLS estimates in Table 1 for comparability with later results.

C.2 Simulation Procedure

While our estimated model characterizes the distribution of income growth rates conditional on innovation, we can simulate from the model to see what our estimated coefficients imply about the evolution of inequality of income levels. To approximate the evolution of income inequality in levels, we use the following procedure.

We begin with an individual's log average residual earnings over the prior 3 years—the same measure which is subtracted off to calculate our growth rate measure—and add to it a randomly generated growth rate from the fitted quantile model. To do this, we interpolate between fitted quantiles to construct a smooth, continuously differentiable quantile function using a flexible parametric approach proposed in Schmidt and Zhu (2016). This approach allows us to efficiently simulate from the estimated conditional quantile model. Specifically, we construct a mapping from a set of 7 conditional quantiles to a smooth density function from $f(y; \mathbf{q}) \equiv f(y; q_1, \ldots, q_7) \colon \mathbb{R} \to \mathbb{R}_+$ by jointly imposing several restrictions:

1. To the left of q_2, $f(y, \mu_l)$ corresponds with a normal density with location and scale parameters chosen to match the conditional quantile restrictions: i.e., its cdf satisfies $\Phi(q_1, \mu_l, \sigma_l) = 0.05$ and $\Phi(q_2, \mu_l, \sigma_l) = 0.1$. Analogously, to the right of q_6, it follows a normal density $\phi(y, \mu_u, \sigma_u)$ which satisfies $\Phi(q_6, \mu_u, \sigma_u) = 0.90$ and $\Phi(q_7, \mu_u, \sigma_u) = 0.95$.

2. Between q_2 and q_6, the CDF is a cubic spline with knots at $\{q_2, ..., q_6\}$ which satisfies the conditional quantile restrictions, is continuous, and has continuous first and second derivatives. These restrictions can be cast as a linear system of equations which has an exact solution given two additional restrictions on the behavior of the spline at the boundaries. We impose that the CDF has a continuous first derivative (i.e., the implied density matches the normal used in the tails).

3. If the spline from part 2 is not strictly monotonic, we linearly interpolate the CDF instead.

Since our CDF is strictly increasing by construction, it is also straightforward to compute the quantile function $Q(u; \mathbf{q})$ by inverting it.

Since the spline is defined using a system of (tridiagonal) linear equations, we can quickly and efficiently solve jointly for the spline parameters for a large number of observations at once. Our simulation exercise is performed as follows:

1. For each individual, using the observed values of income growth y_{it} and covariates x_{it}, compute the CDF of person i's income growth realization: $\hat{u}_{it} \equiv F(y_{it}; \mathbf{q}(\mathbf{x_{it}}, \hat{\beta}))$. $\hat{u}_{it}$ is the percentile of the shock that person i received at time t according to our fitted quantile model.

2. Then, we compute the counterfactual income growth realization as $\tilde{y}_{it} \equiv Q(\hat{u}_{it}; \mathbf{q}(\tilde{\mathbf{x}}_{\mathbf{it}}, \hat{\beta}))$, where $\tilde{x}_{it}$ is a set of counterfactual variables. Note that $\tilde{y}_{it} = y_{it}$ if $\tilde{x}_{it} = x_{it}$.[24]

3. The counterfactual income level is computed by adding the lagged level to the simulated income growth realization $\tilde{y}_{it}$.

The simulated draws associated with the covariates from year t provide an estimate of the cross-sectional distribution of average income from $t + 1$ through $t + 5$ implied by the baseline model when we use the actual values of x_{it} observed in the data.[25] Other than some minor differences in levels introduced by truncation of extreme growth rates in individual earnings[26], our method reproduces the empirical distribution of income levels at each period. The realized change line in Figure 8 is computed by taking the log difference between the set of simulated income statistics in year t relative to the the same statistic for year $t - 5$ times 100, expressed in annualized units (by dividing by 5).

We then compare this realized change with an analogous set of simulations using the fitted quantile functions that obtain when we change the firm and competitor innovation measures, holding all other individual and firm-specific variables fixed. By comparing the simulated level of inequality obtained from these alternative scenarios with the baseline specification, we obtain a model-implied decomposition of the potential contribution of innovation in year t to inequality, relative to this

[24] To minimize the influence of outliers and avoid numerical instabilities related to evaluating inverse functions, we truncate $\hat{u}_{it}$ to the [0.0005,0.9995] interval. Therefore, the simulated draws exhibit slightly less dispersion relative to the original data. We use the same adjustment for both "actual" and counterfactual series, which still allows for an apples-to-apples comparison.

[25] In an earlier draft, we simulated draws from our estimated model for income growth rates by drawing a sample of i.i.d uniform random variables, then evaluate the interpolated quantile function at these points. While results were similar, our current approach has the advantage of maintaining the correlation structure of income realizations across workers at each point in time.

[26] We verify that growth rates associated with the truncated series and raw sample moments are essentially identical.

alternative scenario. Such a calculation is similar in spirit to the decomposition exercises of Machado and Mata (2005) and Firpo, Fortin, and Lemieux (2011), which seek to quantify the predicted effect of changes in the distribution of explanatory variables on quantiles of the unconditional distribution of an outcome variable. While these decompositions do not necessarily have a causal interpretation, they help to shed light on the magnitudes associated with our estimated coefficients.

We compare the simulated levels of earnings inequality implied by the observed distribution of innovation across firms with a simple alternative scenario which replaces $A_{f,t}$ and $A_{I\setminus f,t,t}$ with their equal weighted means across firms within the industry in a given year: $\frac{1}{N_{I(f),t}}\sum_{f\in I(f)} A_{f,t}$ and $\frac{1}{N_{I(f),t}}\sum_{f\in I(f)} A_{I\setminus f,t,t}$, respectively. Whereas the observed distribution of innovation is unequally distributed across firms, this calculation assumes instead that it is symmetric within a given industry-year, but allows for heterogeneity across industries and over time.

We also ran a version of the simulation exercise which replaces $A_{f,t}$ and $A_{I\setminus f,t,t}$ with averages calculated across all firm-years within the same industry for the 1980-1984 period. In this case, in addition to imposing symmetry, we shut down time series variation in the level of innovation within an industry, noting that innovation later in our sample period was generally higher than the observed levels in 1980-1984. Comparing the first and second scenarios speaks to the effects of time-series variation in average industry-level innovation which occurred during our sample period. We find that results are fairly similar to the simulation exercise in the main paper, suggesting the dominant force is actually within-industry heterogeneity in innovation at the same point in time, rather than variation in the level of innovation over time, that drives our main results. These results have been suppressed for brevity.

We compute two simple univariate statistics to summarize the changes in inequality induced by these changes in the innovation measures (results are similar for other percentiles) in the right tail and left tail of the distribution of income, respectively. To capture changes in the right tail, we compute the log difference (times 100) between distance between the 95^th and 50^th percentiles of earnings levels from the baseline and alternative models. Analogously for the left tail, we compare the distances 50^th and 5^th percentiles of earnings levels between baseline and alternative simulations. While we focus on these statistics for brevity, results are similar for other percentiles.

We also decompose the implied changes in inequality into in between and within-firm components. For this analysis, we require that a firm is associated with at least 20 observations in a given year. Our approach follows the construction of Song et al. (2019). For each individual in our sample, we compute the average level of simulated log earnings associated with all of the his coworkers as of time t. We then separately plot the changes in the distribution of average firm log earnings and the distribution of within-firm earnings, which is defined by subtracting the firm-level average from an individual's log earnings. There are two key differences between our construction and the one in Song et al. (2019). First, our earnings measure is cumulated over a longer period of time. Second, firm average earnings are defined relative to the set of coworkers as of time t, which eliminates the need to define a new set of co-workers for purposes of the decomposition.

References

Angrist, J., V. Chernozhukov, and I. Fernández-Val (2006). Quantile regression under misspecification, with an application to the us wage structure. *Econometrica 74*(2), 539–563.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74*(2), 431–497.

Autor, D., D. Dorn, G. H. Hanson, and J. Song (2014). Trade adjustment: Worker-level evidence. *The Quarterly Journal of Economics 129*(4), 1799–1860.

Bena, J. and E. Simintzi (2019). Machines could not compete with chinese labor: Evidence from US firms' innovation. *Working Paper*.

Carley, M., D. Hegde, and A. C. Marco (2014). What is the probability of receiving a U.S. patent? *Yale Journal of Law and Technology forthcoming.*

Firpo, S., N. Fortin, and T. Lemieux (2011). Decomposition methods in economics. *Handbook of labor economics 4*, 1–102.

Guvenen, F., S. Ozkan, and J. Song (2014). The Nature of Countercyclical Income Risk. *Journal of Political Economy 122*(3), 621–660.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2005). Market value and patent citations. *The RAND Journal of Economics 36*(1), pp. 16–38.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Machado, J. A. and J. Mata (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics 20*(4), 445–465.

Rauh, J. D. and I. Stefanescu (2009). Why are firms in the united states abandoning defined benefit plans? *Rotman international journal of pension management 2*(2).

Rauh, J. D., I. Stefanescu, and S. P. Zeldes (2019). Cost saving and the freezing of corporate pension plans. *Working Paper*.

Schmidt, L. D. and Y. Zhu (2016). Quantile spacings: A simple method for the joint estimation of multiple quantiles without crossing.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *The Quarterly Journal of Economics 134*(1), 1–50.