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, Revised June 2021

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, Nicholas Bloom, Carter Braxton, 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, Nikolai Roussanov, Martin Schneider, Toni Whited, and seminar participants in the Social Security Administration, Bank of Portugal, CEAR/GSU, Carnegie Mellon, Duke University, Duke/UNC Asset Pricing Conference, University of Chicago, UNC, the LAEF/Advances in Macro Finance workshop, Minnesota Asset Pricing Conference, Massachusetts Institute of Technology, University of Michigan, Northwestern, Stanford, University of Pennsylvania, 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, Brice Green, Luxi Han, Alejandro Hoyos Suarez, Kyle Kost, Andrei Nagornyi, Bryan Seegmiller, Jinpu Yang, 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. 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.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

Technological Innovation and Labor Income Risk
Leonid Kogan, Dimitris Papanikolaou, Lawrence D. W. Schmidt, and Jae Song
NBER Working Paper No. 26964
April 2020, Revised June 2021
JEL No. E24,G10,G12

ABSTRACT

Using administrative data from the United States, we document novel stylized facts regarding technological innovation and the riskiness of labor income. Higher rates of industry innovation are associated with significant increases in labor earnings for top workers. Decomposing this result, we find that own firm innovation is associated with a modest increase in the mean, but also variance, of worker earnings growth. Innovation by competing firms is related to lower, and more negatively skewed, future earnings. We construct a structural model featuring creative destruction and displacement of human capital that replicates these patterns. In the model, higher rates of innovation by competing firms increases the likelihood that both the worker and the incumbent producer are displaced. By contrast, a higher rate of innovation by the worker's own firm increases profits, but is a mixed blessing for workers, as it increases odds that the skilled worker is no longer a good match to the new technology. Estimating the parameters of the model using indirect inference, we find significant welfare losses and hedging demand against innovation shocks. Consistent with our model, we find that these left tail effects are more pronounced for process improvements, novel innovations, and are concentrated in movers rather than continuing workers.

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

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

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

For most households, human capital is the largest component of total wealth. Unlike financial wealth, human capital is illiquid, poorly diversified, and shocks to its value are largely uninsurable. Thus, understanding what makes human capital risky is important. Our paper focuses on the role of innovation: models of creative destruction generate sharp predictions regarding the rate of technological innovation and redistribution of profits across firms. To the extent that part of a worker's skill set is specific to a particular technology vintage or firm or firms share profits with workers, increased rates of innovation will also affect worker earnings. We explore these ideas in detail and show that there is a strong link between the rate of technological innovation and risk in worker earnings—particularly for workers at the *top* of the earnings distribution, who often have specialized skills and compensation tied to firm profits.

We begin by presenting a set of new stylized facts using millions of administrative worker earnings records from the Social Security Administration (SSA) and the firm-level measure of the value of innovation developed in Kogan, Papanikolaou, Seru, and Stoffman (2017).[1] We first examine industry-level outcomes. We show there is a strong correlation between the rate of technological progress in a given industry and the variation of future firm profits and worker earnings growth. This correlation is significantly stronger for workers at the top of the earnings distribution. The rest of the paper focuses on understanding the drivers of this correlation, and, to that end, for the remainder of the paper we focus on worker-level outcomes.

Empirically, we find that the above relation between innovation and human capital risk is driven by a combination of between- and within-firm effects. We next examine separately the role of innovation that originates in the workers' own firm versus other competing firms in the same industry. We find that increases in the rate of a firm's own innovations are associated with increased future earnings growth for its workers. By contrast, an increased rate of innovation by competing firms is associated with significantly lower future worker earnings. The magnitude of these effects is sizable; the own-firm effects are consistent with most extant estimates of profit sharing elasticities. However, keeping constant the associated change in expected firm profits from new innovations, workers' earnings growth is *more* sensitive to innovation by competing firms relative to own firm innovation. Thus, differences in firm outcomes partly drive workers earnings risk, and workers bear more of that risk on the downside.

[1] 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.

These average effects are informative regarding the link between differences in the rate of innovation across firms and earnings risk. However, they offer an incomplete characterization of effects on income risk: average effects mask considerable heterogeneity in worker outcomes, even for workers whose firms innovate at similar rates. Thus, we next use quantile regressions to characterize how the entire distribution of worker earnings growth rates shifts following innovation by the firm, or its competitors. We uncover a significant relation between the rate of innovation and the higher moments of the distribution of earnings growth. Subsequent to innovation by their own firm, the distribution of earnings growth for the firm's own workers becomes more positively skewed: 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: most workers experience small declines in income, while a minority experiences a significant drop in labor earnings. Importantly, these effects are significantly larger in magnitude for the highest paid workers (top 5% in the distribution of prior earnings within the firm). In addition to larger magnitudes, these top workers also experience a substantial fattening of the left tail of earnings growth when their own firm innovates. For these workers, an increase in innovation of their own firm is a mixed blessing, since it represents a positive shock to both the mean and the variance of their future earnings growth.

In brief, we uncover an economically significant relation between the rate of technological innovation and the earnings risk of top workers. This relation is striking given that the existing literature has largely emphasized the complementarity between technology and high-skill workers (Goldin and Katz, 1998). To help interpret our facts, we develop a model of innovation and earnings risk. Our model focuses on high-income (i.e. skilled) workers and has the following key ingredients. Firms compete in the product market; similar to quality-ladder models, only the most efficient firm finds it profitable to produce in a given product line. Firm profits (markups) depend on the distance between the leading producer (the incumbent) and the next most efficient firm (the potential entrant). High rates of innovation by the entrant result in lower profits by the incumbent firm and potentially the loss of the market for that product line. Firms hire skilled workers to manage these product lines and these managers' earnings are exposed to profits to mitigate a moral hazard friction; hence, there is pass-through of firm profits to the earnings of top workers. If the firm loses its technology lead in a product line, the incumbent manager is displaced. Innovation by the worker's own firm leads to higher profits, but the compatibility of these innovations with the skills of the incumbent worker is uncertain. That is, the productivity

of the incumbent worker in utilizing the new vintage has a stochastic component and may increase or decrease; in some cases, the firm may find it profitable to replace the incumbent worker with a new hire.

In sum, the model generates a link between innovation, firm profits, and worker earnings risk. In the model, innovation by competing firms increases the likelihood that both worker and incumbent producer are displaced. Thus, it is associated with lower future firm profits and a more negatively skewed distribution for earnings growth. By contrast, higher rates of own firm innovation lead to higher average profits but are a mixed blessing for (top) workers, since they increase the likelihood the manager is no longer a good match to the firm's technology. Hence, the model generates a higher mean, but also an increase in the variance of earnings growth of top workers in response to innovation by their own firm. Exposures are also asymmetric: managers capture only a fraction of the gains from own firm innovation but fully share in losses from competitor innovation.

We calibrate the model to match our stylized facts above using indirect inference. The model can quantitatively replicate the key relation between technological progress and earnings risk that we uncover. In addition, the calibrated model allows us to quantify the impact of innovation on worker utility. We find that innovation is associated with significant welfare losses for top workers, even in the presence of progressive taxation and ability to self-insure by accumulating assets. In the model, the average worker would need to receive a proportional subsidy of 1.5% in her lifetime consumption to offset the utility loss resulting from a one-standard-deviation (one time, but persistent) increase in industry innovation. Our estimates thus imply a substantially higher welfare cost of innovation for (top) workers than the welfare cost of business cycles due to job displacement (Krebs, 2007).

In addition, we use the model to compute the willingness of workers to invest in assets that (partially) hedge their earnings risk. We find significant demand for insuring changes in the rates of innovation in the worker's own industry, as measured by fluctuations in their marginal utility. Specifically, a one standard deviation increase in industry innovation leads to a 0.38 log point increase (on average) in the marginal utility of skilled workers. To a first approximation, this estimate implies that workers would be willing to invest in an asset whose return is maximally correlated with a shock to industry innovation as long as its Sharpe ratio (market price of risk) was higher than -0.38. That said, there is dispersion across model workers in the magnitude of these welfare costs and their willingness to hedge. Similar to the data, higher-income workers in the model face higher risk exposures and are therefore willing to pay more for insurance.

Our model has testable predictions that we examine in our data. First, the model generates an increased left tail of earnings growth through a higher likelihood of job loss. An advantage of our data is that they allow us to track workers across firms and therefore to examine the extent to which this increase in the risk of earnings declines is related to separations. We therefore explore how the magnitude of the increase in the left tail varies between stayers (continuing workers) and movers (workers who leave the firm). In both the model and the data, an increase in the rate of innovation by either the firm or its competitors is associated with a significantly higher increase in the left tail among movers than stayers. Further, and consistent with displacement of human capital, incumbent workers are more likely to experience persistent unemployment spells. As before, these effects are larger in magnitude for top workers.

Second, our model has specific predictions about what types of innovations are more likely to lead to higher worker earnings risk. In particular, we expect the effects of own firm innovation to be driven by process (as opposed to product) improvements, and by innovations that are significantly different from what the firm has done in the past. We measure the former using the process/product characterization of patents of Bena and Simintzi (2019). We measure novelty using the methodology of Kelly, Papanikolaou, Seru, and Taddy (2020), who quantify textual similarity across patent documents. A patent is novel if the description of the innovation is sufficiently distinct from the firm's prior innovations. Consistent with our model, we find that the increase in the left tail of (top) worker earnings is driven by process improvements. By contrast, non-process (product) improvements are primarily associated with an increase in average earnings rather than an increase in higher moments. Further, the magnitudes are significantly stronger for innovations that are novel to the firm; the impact of non-novel innovations for the left tail is essentially zero.

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.[2] 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

[2]Some of the instruments used in the literature for the granting of a patent do not apply in our setting. Specifically, Sampat and Williams (2019) use the random assignment of a patent to examiners with different propensities to approve a patent application to instrument for patent grants. However, our sample focuses on large, publicly traded firms, many of which file hundreds, if not thousands, of patent applications in a given year. To the extent that assignment were indeed random and independent of firm characteristics, we would expect these examiners fixed effects to be diversified away.

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, 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.

Overall, we provide a set of novel stylized facts regarding the relation between innovation and worker earnings risk; we interpret these facts through the lens of a structural model of firm innovation and earnings risk and provide evidence that the model's testable implications are supported by the data. Our main conclusion is that innovation is associated with increased earnings risk for top workers and that this mechanism operates through a combination of profit-sharing and skill displacement. As such, our paper is connected to several strands of the literature.

Our focus on the earnings risk of top workers distinguishes our work from most of the existing work studying the link between technological innovation and worker earnings. 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; Adão, Beraja, and Pandalai-Nayar, 2020); or the substitution between workers and new forms of capital (Hornstein, Krusell, and Violante, 2005, 2007; Acemoglu and Restrepo, 2020). Our model combines elements of models with creative destruction (Aghion and Howitt, 1992; Grossman and Helpman, 1991; Klette and Kortum, 2004) and vintage-specific human capital (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002). Our findings are particularly striking in light of the traditional view that technology tends to complement high-skill labor (Goldin and Katz, 1998; Krusell, Ohanian, Ríos-Rull, and Violante, 2000; Goldin and Katz, 2008; Atalay, Phongthiengtham, Sotelo, and Tannenbaum, 2020). A likely source of this difference is our focus on patents by individual firms, which implies that we necessarily study innovation rather than adoption of existing technologies (for instance, robots or automation in general). More broadly, however, it is likely that process improvements are often associated with significant organizational changes, which may lead to the replacement of mid-level executives that lack the skills, or willingness,

to adapt to new production methods (Davenport, 1993).[3] As such, our findings complement the findings of Deming and Noray (2020), specifically, that individuals in occupations with greater changes in skill requirements have lower returns to experience (possibly due to faster skill obsolescence).

A key part of our model mechanism operates through job loss. As such, our work connects to the literature studying earnings losses of displaced workers (Gibbons and Katz, 1991; Neal, 1995; Huckfeldt, 2018; Jarosch, 2021; Braxton and Taska, 2020). Closest to our paper is Braxton and Taska (2020), who show that individuals displaced from occupations undergoing greater amount of technological change experience larger earning declines following job loss. Our finding that new innovations are associated with substantial displacement risk for high income workers is consistent with the mechanism in the Jones and Kim (2018) model of top income inequality. They also relate broadly to papers which have used spatial variation to link innovation, top income inequality, and social mobility (Aghion, Akcigit, Bergeaud, Blundell, and Hémous, 2019; Aghion, Akcigit, Deaton, and Roulet, 2016; Akcigit, Grigsby, and Nicholas, 2017).

Our focus on the interaction between technological progress, product market competition, skill displacement and worker earnings *risk* sharply differentiates our work from studies that examine 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.[4] Contributing to this literature, our quantile regressions reveal substantial heterogeneity in worker outcomes following technological improvements by the firm—or its competitors—that are otherwise obscured when focusing on average (i.e. conditional mean) outcomes. This is particularly important, in light of the fact that the existing literature has

[3] As an example, Davenport (1993) discusses 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. [...] 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, pp.168–170, 194–195).

[4] Card, Cardoso, Heining, and Kline (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. A related literature has considered earnings responses of inventors, firm owners, and CEOs (Toivanen and Väänänen, 2012; Frydman and Papanikolaou, 2018; Bell, Chetty, Jaravel, Petkova, and Van Reenen, 2019; Akcigit et al., 2017; Aghion, Akcigit, Hyytinen, and Toivanen, 2018) to own firm innovation.

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; Ellul, Pagano, and Schivardi, 2017). We view our work as complementary; rather than focusing on estimating profit-sharing elasticities, our goal is to quantify the relation between innovation and the uncertainty in worker earnings, understand its determinants, and explore its implications. To that end, we view our structural model as a valuable tool. We also point to how displacement effects linked with shifts in technology likely limit potential for risk sharing through the firm.

Our finding that innovation is associated with an increase in the left tail of earnings growth for top earners makes it increasingly likely that it matters for asset prices, especially given the high concentration of stock ownership (and participation) among the richest households (Poterba and Samwick, 1995). In particular, our empirical estimates and structural model imply that top workers experience a significant increase in marginal utility in response to high degrees of technological innovation in their own industry. This increased demand for insurance against states with high degrees of technological innovation contributes to a negative risk premium for (displacive) technology shocks, reinforcing the implications of Papanikolaou (2011); Garleanu, Kogan, and Panageas (2012); Kogan, Papanikolaou, and Stoffman (2020) for the equity premium and the value spread. In these models, investors are willing to invest in innovative (i.e. growth) firms, despite their lower than average returns, because their returns are positively correlated with households' marginal utility.[5] We contribute to this literature by providing direct evidence that technological innovation is correlated with worker earnings risk. Our findings also suggest that allowing for displacement of human capital in the model of Kogan et al. (2020) reinforces its main mechanism and would allow it to match the observed properties of value and growth firms with only moderate levels of risk aversion. More broadly, our findings connect to a recent literature arguing for the cyclical properties of skewness in labor income and its implications for the equity premium (Guvenen, Ozkan, and Song, 2014; Constantinides and Ghosh, 2017; Schmidt, 2016). In terms of magnitudes, the increase in the left tail of (top) worker earnings 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).

[5]Over the last decade, growth firms (as traditionally classified) have not exhibited significantly different returns than value firms. Though estimating expected returns by looking at average returns over a relatively short period is fraught with pitfalls—especially in the presence of positive surprises to innovation outcomes—another important factor is related to mis-measurement of book values (the value of assets in place) as it typically omits intangible assets, which have arguably become more important over time (Park, 2019; Eisfeldt, Kim, and Papanikolaou, 2020).

Last, our work also connects 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, Price, Guvenen, Bloom, and Von Wachter (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. To relate our findings to this literature, in the appendix we perform a simulation-based decomposition exercise which quantifies implication of our empirical estimates for the recent rise in income inequality among firms. During the 1990s, both the level as well as the dispersion in innovation outcomes across firms increased; 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.

1 Data and Measurement

We begin by briefly summarizing the data on labor income and firm innovation outcomes used in our analysis. All details are relegated to Appendix A.

1.1 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). 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.[6] We follow Guvenen et al. (2014) and exclude self-employed workers and individuals with earnings below a minimum threshold—equal to the amount one would earn working 20 hours per week for 13 weeks at the federal minimum wage. See Appendix A.1 and

[6]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. Our sample is the same as Guvenen et al. (2014). 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.

Guvenen et al. (2014) for further details.

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

$$w^i_{t,t+h} \equiv \log\left(\frac{\sum_{j=0}^{h} \text{W-2 earnings}_{i,t+j}}{\sum_{j=0}^{k} D(\text{age}_{i,t+j})}\right). \tag{1}$$

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 (1) 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 (1) 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 (1) over a horizon of h years:

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

For the bulk of our analysis we will focus on 5 year horizons, $h = 5$. Examining (2), 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 (2) 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 A.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. Importantly, since we can track workers across firms, their earnings growth rate (2) may include income from more than one employer.

1.2 Innovation Outcomes

Our main independent variables of interest capture the rate of innovation 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 each worker via the firms with whom she is employed, 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. That is, we will *not* be measuring firms' adoption of technologies developed by other firms. Therefore, 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 (2020)).

A major challenge in measuring innovation by using patents is that patents 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

$$A_{f,t} = \frac{\sum_{j \in P_{f,t}} \xi_j}{K_{ft}} \tag{3}$$

and its competitors,

$$A_{I \setminus f,t} = \frac{\sum_{f' \in I \setminus f} \left(\sum_{j \in P_{f',t}} \xi_j \right)}{\sum_{f' \in I \setminus f} K_{f't}}. \tag{4}$$

In addition, we can aggregate the above to the level of the industry

$$A_{I,t} = \frac{\sum_{f' \in I} \left(\sum_{j \in P_{f',t}} \xi_j \right)}{\sum_{f' \in I} K_{f't}}. \tag{5}$$

Here, ξ_j corresponds to the KPSS value of a patent, and $P_{f,t}$ denotes the set of all patents used to measure innovation by firm f during period t. 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 the firm's size K. We use the firm's capital stock (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.9). Appendix A.4 provides more details on the construction of these variables. In the context of the model in Section 3, we can interpret $A_{f,t}$ and

$A_{I\setminus f,t}$ as empirical proxies for the rates of innovation $\lambda_{f,t}$ and $\lambda_{f',t}$.[7]

A potential shortcoming of a patent-based measure 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 it becomes known that the patent application has been successful, which 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 when applications for these patents are filed. Hence, when computing $A_{f,t}$, the set of patents $P_{f,t}$ includes patents that are *filed* in year t.[8] Consistent with this timing convention, Appendix Figure 2 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 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. We also use the issue date when constructing $A_{I,t}$. That said, this choice of timing is not the 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.

1.3 Overview of the sample

Our final matched sample includes approximately 14.6 million worker-year 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 public firms' innovation outcomes with individual workers' earnings histories using EIN numbers. A worker is included in the sample in year t if she works in a matched

[7] 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.

[8] 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).

firm. But, since workers can transition to other firms potentially not in our sample, our calculation of future earnings growth rates (2) includes her earnings in any new firm she (possibly) transitions to.

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 and Figure A.1 for additional details). The industry composition of the matched sample (to Compustat firms) and the unmatched workers 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.

2 Technological Innovation and Risk

Here, we document a number of novel stylized facts regarding the relation between innovation and earnings risk.

2.1 Industry Innovation and Risk of Firms and Workers

We begin by documenting the correlation between technological innovation in a given industry and the variability of workers' labor earnings and firm profits. We measure risk as the cross-sectional dispersion in worker earnings and firm profits. We measure worker risk as the variance in worker earnings growth $g_{i,t:t+h}$, defined in (2), over a horizon of h years:

$$v_{I,a,w}^{t:t+h} = \sum_{i \in (I,w,a)} \left(g_{i,t:t+h} - \bar{g}_{i,t:t+h}\right)^2. \tag{6}$$

To obtain an estimate of earnings worker risk, we condition on observable characteristics. As a result, our measure of worker labor income risk varies by calendar year t; industry I, defined at the 3-digit SIC level; worker age a, grouped into 10 bins; and past earnings level w, grouped into bins and based on $w_{t-4,t}^{i}$ within each age-industry-year cell. Accordingly, we construct a measure of profit dispersion across firms within an industry,

$$v_{I}^{t:t+h} = \sum_{f \in I} \left(g_{f,t:t+h} - \bar{g}_{f,t:t+h}\right)^2. \tag{7}$$

Here, firm profit growth is defined as the growth in cumulative gross profits Π, analogously to equation (2),

$$g_{f,t:t+h} \equiv \log \left[\frac{1}{|h|} \sum_{\tau=1}^{h} \Pi_{f,t+\tau} \right] - \log \Pi_{f,t}. \tag{8}$$

The definition of gross profits is equal to revenue minus costs of goods sold. Similar to (2), our focus on cumulative profits emphasizes persistent profit growth and helps smooth transitory shocks in profitability.

We next estimate the relation between technological innovation at the industry level $A_{I,t}$, defined in equation (5), and the cross-sectional dispersion in firm profits and worker earnings growth and innovation in the industry,

$$\log v_I^{t:t+h} = \beta \, A_{I,t} + \rho \, \log v_I^{t-5:t} + c_t + u_{I,a,w,t} \tag{9}$$

and

$$\log v_{I,a,w}^{t:t+h} = \beta \, A_{I,t} + \rho \, \log \, v_{I,a,w}^{t-5:t} + c \, Z_{I,a,w,t} + u_{I,a,w,t}. \tag{10}$$

Our coefficient of interest is β, which captures the association between earnings risk and industry innovation $A_{I,t}$. In the case of worker earnings, we allow the sensitivity to vary by the level of worker earnings. The vector of controls Z includes a battery of fixed effects and interactions: industry-age; age-income; year-age; and year-income fixed effects. We normalize A_I to unit standard deviation, and cluster the standard errors by industry, worker age and past income. Figure 1 presents the estimated coefficients β for horizons of $h = 3$ and $h = 5$ years.

Overall, we note a significant association between the level of technological progress in a given industry and dispersion in future firm profitability. A one standard deviation increase in A_I is associated with approximately a 0.075 log point increase in the cross-sectional dispersion of profits over the next 3 or 5 years. The fact that results are highly comparable across horizons suggests that this is a highly persistent increase in firm-level risk.

Moreover, the right side of the figure shows that an increase in innovation is associated with a significant increase in the variability of worker earnings over the same horizon—but only for the workers at the top of the income distribution. Specifically, a one standard deviation increase in A_I is associated with a 0.05 and 0.095 log point increase in worker earnings risk for workers in the 75^{th} to 95^{th} and above the 95^{th} percentile, respectively. As before, these magnitudes do not vary materially with the horizon over which we measure earnings, implying that these are highly persistent changes.

In brief, we see that increased innovation at the industry level is associated with increased dispersion in both firm profits and earnings of top workers. Schumpeterian growth models imply that technological innovation is typically associated with substantial reallocation and creative destruction, leading to winners and losers in the cross-section of firms. Kogan et al. (2017) provide a battery of evidence consistent with this prediction. To the extent that firms share some of their profits with (a subset of) workers, we would also expect to see a similar correlation between innovation and uncertainty in labor income for these workers. In the next two sections, we explore this idea in more detail using firm and worker-level data.

2.2 Innovation, Firm Profits, and Worker Earnings

To understand why industry innovation is associated with increases in risk for both firms and (top) workers, we analyze the data at a higher level of granularity by examining outcomes for firms and individual workers. Doing so also allows us to estimate differential effects depending on where the innovation occurs, that is, we can separate innovation by the firm versus its competitors. Kogan et al. (2017) show that differences in innovation outcomes are associated with substantial heterogeneity in subsequent growth in profitability.

We begin by estimating a slightly modified specification than KPSS, in which the dependent variable is the growth rate in cumulative profits (8), in direct analogy to our worker earnings growth measure (2),

$$g_{f,t:t+h} = a\, A_{f,t} + b\, A_{I\setminus f,t} + c\, Z_{ft} + u_{ft}. \tag{11}$$

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). 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 $A_{I\setminus f}$ to unit standard deviation.

In addition, we estimate the response of worker earnings growth using a similar specification as above

$$g_{i,t:t+h} = a\, A^{sm}_{f,t} + b\, A^{sm}_{I\setminus f,t} + c\, Z_{i,t} + \varepsilon_{i,t}, \tag{12}$$

The vector of controls Z includes the same set of firm-level controls as (11); to that, we add a battery of worker-level 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 rates.[9] 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 (the analogue of clustering by firm). See Appendix C for more details.

Our main coefficients of interest are a_h and b_h, which measure the response in firm profits and worker earnings to innovation by the firm and its competitors, respectively. Panel A of Table 1 presents our estimates for horizons $h = 5$ 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. Figure 2 presents results across horizons; as we compare the estimates between horizons of five to ten years, we see that these are largely permanent effects.[10]

Panel B of Table 1 reports our estimates for the response of worker earnings. Focusing on all workers, 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

[9] 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, and we allow these coefficients to differ across five bins formed based upon a worker's rank within the firm (discussed in footnote 11).

[10] Figure 2 plots the estimated coefficients a_h and b_h for values of $h = -5$ to $h = 10$. This allows us also to examine whether innovation is related to past trends in profitability. That is, one potential concern is that firm innovation is related to some unobservable source of heterogeneity that itself is responsible for increased firm profits. 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—which lends support to our convention for dating patents. 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.

innovate. Appendix Table A.3 presents results for additional horizons; as we compare coefficients horizons, 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 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 0.08 log point increase in profitability compared to a 0.014 log point increase in earnings for the firm's own workers. These numbers imply a profit-sharing elasticity approximately equal to $1.4/8 \approx 0.17$ for all workers and 0.195 for stayers. To put these numbers in context, 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, respectively.

Importantly, however, we note that the profit sharing elasticity implied by 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 we outline below: 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.

The right panel of Table 1 examines how this sensitivity varies by the worker's earnings percentile relative to other workers in the same firm.[11] Overall, we note that the sensitivity of worker earnings to firm or competitor innovation is generally higher for the top workers. For example, a one standard deviation increase in A_f is associated with a cumulative increase of 1.53% to 1.85% in earnings for workers in the top quartile, compared to 1.19% for the workers in the bottom quartile. This difference

[11] We follow Guvenen et al. (2014) and compute worker earnings ranks based on the last 5 years of earnings—that is, we sort workers by $w_{t-4,t}$, defined in equation (1) within the firm. Whenever we allow a and b 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. Here, we 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. As an additional 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 qualitatively very similar (see Appendix Figures A.11 to A.13 for details).

in sensitivity is much more stark when we examine responses to competitor innovation $A_{I\setminus f}$. A one-standard increase in $A_{I\setminus f}$ is associated with a 2.2% to 5.9% decline in earnings for workers in the top quartile, compared to a 1.9% decline for workers in the bottom quartile.

Our findings so far do not differentiate among workers in the same firm. That is, they correspond to the conditional mean of earnings growth faced by a particular worker employed by a given firm, where either the firm itself or its competitors innovate. More broadly, however, innovation may affect not only the conditional mean, but also the conditional variance—or higher moments—of earnings growth. Thus, focusing on average responses can mask substantial heterogeneity in ex-post outcomes across workers.

We next examine how the *conditional distribution* of worker earnings growth rates is related to innovation by workers' employers or their competitors. In particular, we next estimate the response of individual quantiles in worker growth rates $g_{i,t:t+h}$ using a specification analogous to equation (12). The only difference is that now, 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 separability restrictions and methodology for jointly estimating multiple conditional quantiles of Schmidt and Zhu (2016), who assume that the median and log of the difference between each two adjacent quantiles follow a linear model. 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. Figure 3 plots the average marginal effects of a one standard deviation change in each variable of interest on each conditional quantile of the earnings growth distribution. The top row presents results for all workers, whereas the bottom row allows the response coefficients to firm (a_h) and competitor (b_h) innovation to vary with the worker's current earnings rank.

Examining Figure 3, two patterns stand out. First, the shift in average worker earnings we documented in Table 1 is distributed asymmetrically across workers. In particular, innovation is associated with shifts in both the variance and the skewness of future worker earnings growth. Second, the magnitudes are significantly larger for top workers. We next discuss each in more detail.

Focusing on all workers employed by innovating firms (Panel A), 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, 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, 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.

Panel B of Figure 3 examines 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.

The bottom row of Figure 3 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 substantially larger increase in dispersion (and skewness) in response to innovation than the income growth rates of lower-paid workers. We see that workers in the top 5% and bottom 95% experience qualitatively similar increases in skewness in income growth 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%.

Importantly, we also see in Panel C that 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 and large income drops. As a result, the impact of own firm innovation on the utility of top workers is theoretically ambiguous and

depends on their risk aversion.

Last, Panel D shows that top workers are also more likely to experience large income drops following higher innovation output by competing firms. Specifically, 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. 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.

2.3 Summary and Discussion

The stylized facts in the previous section can be summarized as follows. We see that own-firm innovation is followed by higher firm profits and (mostly) improved outcomes for the firm's workers. By contrast, innovation by competing firms is associated with lower profits for the focal firm and unambiguously worse future outcomes for its workers. In addition, these mean effects mask significant heterogeneity in outcomes: from the perspective of an individual worker, innovation is associated not only with shifts in mean earnings growth, but also changes in both the variance and the skewness of future worker earnings growth. These effects are larger in magnitude for top workers and are comparable in magnitude across horizons, suggesting that they represent permanent shifts.

Our preferred interpretation of these facts is that innovation is associated with an increase in the riskiness of labor income, particularly for the firms top workers. This interpretation does not automatically follow from these results; it is possible that the heterogeneity in worker outcomes we observe following the innovation shocks is perfectly predictable from the perspective of an individual worker, who has more knowledge into her labor income process than the econometrician. That said, it is useful to keep in mind that we are comparing workers that are similar in terms of ex-ante observable characteristics—our specifications include a rich set of control variables—which include flexible functions in worker age, and the level and growth rate in past earnings.

3 The Model

In this section we present a structural model that generates a link between technological progress in a given industry and riskiness of labor income of incumbent workers. The model combines elements of

models with creative destruction (Aghion and Howitt, 1992; Grossman and Helpman, 1991) and vintage-specific human capital (Chari and Hopenhayn, 1991; Jovanovic and Nyarko, 1996; Violante, 2002).

Our model features two key mechanisms through which innovation increases labor income risk. First, firms share profits with workers. As firms experience increased or decreased profits in response to successful innovation, the earnings of top workers respond similarly—consistent with our findings in Table 1. Second, part of a (skilled) worker's human capital is specific to a particular vintage of technology. As firms improve on their production process, they may discover that incumbent workers are not as skilled in the new technology—to the point that the firm may find it optimal to replace them. This mechanism can generate an increase in the left tail of earnings growth for a top worker when her own firm innovates—consistent with Panel C of Figure 3. Further technical details are in Appendix B.

3.1 Setup

A number of firms compete in producing intermediate goods. Output X_t is produced by a continuum of intermediate goods according to

$$\log X_t = \int_0^1 \log x_{i,t} di. \tag{13}$$

Firms can produce each good using a constant returns to scale technology. Firms differ in their level of efficiency in producing each good: good i is produced using the following technology,

$$x_{i,t} = q_{i,t}\, l_{i,t}\, e_{i,t}\, z_{i,t}, \tag{14}$$

where $q_{i,t}$ is the quality of the leading producer and $l_{i,t}$ is a factor of production (unskilled labor or land) in fixed supply (normalized to one), which can be freely reallocated across product lines, and which firms hire at the prevailing equilibrium price. Given our constant returns to scale assumption in (14), only the leading producer (the firm with the highest $q_{i,t}$) finds it profitable to produce each good.

Importantly, the production of an intermediate good requires a skilled worker (manager). Managers differ in their ability: $z_{i,t}$ denotes the ability of the current manager assigned to good i. There is a moral hazard friction, in that the skilled worker can potentially divert output—here, $e_{i,t}$ denotes the fraction of un-diverted output. Diversion is costly: if the manager diverts one 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, $e_{i,t} = 1$.

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

$$\Pi_{i,t}\, dt = \left(1 - \frac{\bar{z}}{\kappa_{i,t}\, z_{i,t}}\right) X_t\, dt. \tag{15}$$

Markups depend on the firm's current technology lead over its closest competitor, which has productivity $\tilde{q}_{i,t} \equiv q_{i,t}/\kappa_{i,t}$, and the productivity of the current manager $z_{i,t}$ relative to the next best alternative a new entrant could hire, $\bar{z}$. We set $\bar{z} = 1$. Due to the moral hazard friction, the skilled worker j assigned to good i receives a fraction of the firm's profits,

$$w_{j,t}\, dt = \beta\, \Pi_{i,t}\, dt = \beta \left(1 - \frac{\bar{z}}{\kappa_{i,t}\, z_{i,t}}\right) X_t\, dt. \tag{16}$$

The goal of the model is to illustrate the implications of firm innovation (and competitor innovation) for the distribution of worker earnings growth. Given that our empirical specifications include time fixed effects, we can focus our attention on top worker earnings scaled by output, $w_{j,t} = W_{j,t}/X_t$. Next, we need to characterize the evolution of $\kappa_{i,t}$ and $z_{i,t}$ and introduce firms.

A firm is a collection of product lines that it finds profitable to produce—that is, goods in which the firm is the leading producer. Given that we are interested in the workers' earnings process, the exact number of firms competing in each product line does not matter: what matters is the distance between the leading producer and the closest competitor $\kappa_{i,t}$, and the chance that one of the competitors overtakes the incumbent producer. Firm innovation is exogenous and takes the form of improvements in efficiency: over an instant dt, a firm can innovate with flow probability $\lambda_{f,t}\, dt$.

We next map firm (A_f) and competitor ($A_{I\setminus f}$) innovation into the model. Both empirical measures are continuous, and are meant to capture changes in the intensity of a firm's innovation outcomes, as they are aggregated over hundreds or thousands of patents in a given year. In the model, these measures correspond to the innovation arrival intensity $\lambda_{f,t}$, which is a stochastic process. To simplify the analysis, we specify that $\lambda_{f,t} \in (\lambda_L, \lambda_H)$ is a two-point Markov process, with $\lambda_H > \lambda_L$ and generator matrix given by

$$T = \begin{pmatrix} -\mu_H & \mu_H \\ \mu_L & -\mu_L \end{pmatrix}. \tag{17}$$

We further assume that there are two firms competing in the production of each good i. Hence, A_f

corresponds to an increase in λ_f and $A_{I\setminus f}$ corresponds to an increase in $\lambda_{f'}$.

In brief, innovation coincides with the arrival of a Poisson process with intensity $\lambda_{f,t}\,dt$. Conditional on successfully innovating on its own product line, an incumbent firm increases its current technology lead $\kappa_{i,t}$ by a proportional amount Δ_κ. By contrast, innovation by potential entrants can take two forms: radical or incremental. An incremental innovation by the potential entrant leads to a decrease in the incumbent's current technology lead $\kappa_{i,t}$ by a proportional amount h. Radical innovation by the potential entrant results in the firm now becoming the leading producer, drawing a new value for $\kappa_{i,t} \sim f(\kappa; \kappa_{i,t-}, z_{i,t-})$. In order for the new entrant to find it profitable to produce with a new manager with ability $\bar{z} = 1$, we require that the new draw satisfies $\kappa_{i,t} \geq \kappa_{i,t-}\, z_{i,t-}$ by drawing $\kappa_{i,t}$ from a truncated normal distribution following radical innovation. When the competitor is in the low innovation state $\lambda_{f't} = \lambda_L$, all innovations are incremental and occur at rate $\lambda_L\,dt$. In the high innovation state, $\lambda_{f't} = \lambda_H$, radical innovations occur with probability $p\,\lambda_H\,dt$ and incremental innovations occur with probability $(1-p)\,\lambda_H\,dt$.

Skilled workers vary in their productivity z_j. An important assumption of our model is that a skilled worker's productivity z_j is specific to a given technology-firm combination. As such, firm innovation has also a direct effect on worker productivity. That is, if a firm improves on the technology in producing one of its own goods, the skilled worker may not be as productive using the new technology. Conditional on the firm innovating at time t, the incumbent worker's productivity using the new technology vintage is given by a new draw $z_{j,t}$, which is correlated to her productivity using the older vintage $z_{j,t-}$, but also has a stochastic component $\varepsilon_{j,t}$

$$\log z_{j,t} = \mu_z + \rho_z \log z_{j,t-} + \varepsilon_{j,t}, \qquad \varepsilon_{j,t} \sim N(0, \sigma_z^2). \tag{18}$$

Innovation can potentially displace the firms' existing skilled workers. If the existing worker's productivity using the new technology is sufficiently low, the firm may find it optimal to replace the worker with an unemployed manager that has productivity level $\bar{z}$.[12] If the incumbent firm f loses the leading efficiency to a competitor f', the position is eliminated and the worker previously assigned to that good becomes unemployed; the new entrant hires a top worker with ability $\bar{z}$.

Unemployment is transitory. Unemployed workers find a new job with flow probability $\theta\,dt$, in which

[12] Here, it is not crucial that all new hires have the same level of productivity. The decision rule would be the same if productivity of new hires was unobservable by the firm, in which case $\bar{z}$ would correspond to their prior belief about the productivity of new hires.

case they are randomly matched with a firm seeking to hire a new manager. Newly-employed managers start with ability $\bar{z}$. Unemployment is costly, not only due to lost earnings but also because re-hired managers likely start at a lower wage (they fall off the ladder).[13] That said, conditional on being unemployed, all skilled workers are identical, due to our assumption that all skills are firm-specific. This choice reflects our preference for parsimony; allowing for general human capital, which depreciates with the time spent out of employment as in the models of Huckfeldt (2018) and Jarosch (2021) is a potentially interesting extension. While it is straightforward to extend the model to allow workers to have finite working lives by assuming that workers retire with flow probability $\delta\, dt$ per period, we abstract away from this in the current calibration for simplicity.

Overall, we have presented a model where innovation affects the earnings risk of top workers. In particular, we interpret the skilled workers in the model as corresponding to the workers at the top of the earnings distribution in the data. Accordingly, when matching the model to the data, we will focus on matching the properties of income growth for the top 5% of the workers. That said, there are two points worth discussing. First, the model can also generate differences in labor income risk among skilled workers; we briefly discuss this in Section 3.3. Second, the model does feature unskilled workers, but their labor earnings are not risky: unskilled workers are always employed, supplying an effective quantity of labor l at the equilibrium wage. As such, we view the model as providing a useful distinction between the riskiness of labor income for skilled versus unskilled workers, even as our goal is not to match the earnings dynamics of unskilled workers in the data.

3.2 Calibration and Model fit

The model has a total of 10 parameters, which we calibrate using indirect inference (Ingram and Lee, 1991). Specifically, given a vector X of target statistics in the data, we obtain parameter estimates Θ by minimizing the distance between the model and the data,

$$\hat{\Theta} = \arg\min_{\Theta} \left(X - \hat{X}(\Theta)\right)' W \left(X - \hat{X}(\Theta)\right), \tag{19}$$

where $\hat{X}(\Theta)$ is the vector of statistics generated by the model. Our choice of weighting matrix $W = diag(XX')^{-1}$ penalizes proportional deviations of the model statistics from their empirical

[13] We assume that top workers cannot search for a better match while employed. Under this assumption, we verify that in equilibrium it is never optimal for the top worker to separate from the firm voluntarily in hope of later making a better match with another firm. This justifies our treatment of all separations as initiated by the firm.

counterparts. Since the model is aimed to understand the earnings dynamics of top earners, we target statistics of earnings growth for top workers (top 5% rank). Specifically, our calibration targets include the unconditional earnings growth percentiles (Appendix Figure A.4), as well as their response to firm and competitor innovation shocks (Figure 3). Panel A of Table 2 presents the estimated model parameters that pertain to the dynamics of worker earnings.

Figure 4 plots the model fit in terms of the statistics that we target. Panel A compares the unconditional percentiles of earnings growth between the top workers in the data versus the model. Overall, we see that the model performs quite well, as it generates a realistic dispersion in earnings growth rates. Panel B compares the response in the model and the data to an innovation shock, either by the own firm (Panel B.i) or competitors (Panel B.ii). Specifically, Panel B.i compares the response of the mean and percentiles of earnings growth to a one standard deviation increase firm innovation: A_f in the data versus an increase in the firm's innovation rate from $\lambda_f = \lambda_L \to \lambda_H$, scaled so that it corresponds to a one standard deviation increase in λ—which is equal to $\sqrt{\mu_H \mu_L}/(\mu_H + \mu_L)$. Similarly, Panel B.ii compares the response to competitor innovation between the data and the model, $A_{I\setminus f}$ and $\lambda_{f'} = \lambda_L \to \lambda_H$, respectively. To these figures we also add mean earnings growth in the data and in the model, even though means were not explicitly part of the calibration targets. Examining Panel B of Figure 4, we see that the model can qualitatively, and quantitatively, replicate the key stylized features of the data we documented in Section 2.2. Specifically, an increase in the rate of firm innovation increases both the mean and the variance of earnings growth for top workers. By contrast, an increase in competitor innovation is associated with a sizeable increase in the left tail of earnings growth and a decrease in mean earnings growth.

For the remainder of the paper, we focus on aspects of the data that were not explicitly targeted in our calibration. In particular, we next examine the extent to which our model can replicate the stylized facts of Section 2.1, that is, the correlation between earnings risk and innovation outcomes for top workers (top 95%) at the industry level. We can think of industry in the model as referring to a single product line. As such, the model version of an industry shock is an increase in the *average* λ—that is, a simultaneous increase in λ_f and $\lambda_{f'}$ from λ_L to λ_H, appropriately scaled to reflect the standard deviation of the average λ assuming the two shocks are independent. In addition to the (log) increase in the cross-sectional variance of earnings growth, we also report the changes in the individual percentiles in both the model and the data. To obtain the data equivalents, we re-estimate equation (10), replacing the dependent variable with the percentiles of earnings growth, which again vary by industry and worker age.

Figure 5 compares the model to the data. Overall, we see that the model performs reasonably well in generating an increase in worker earnings risk in response to an increase in industry innovation. For example, a one standard deviation increase in the level of industry innovation is followed by a 0.125 log point increase in the variance of worker earnings growth over the next 5 years, compared to 0.095 log points in the data; the difference falls within two standard errors of the empirical estimate. To appreciate this result, note that even though we did target the response of the moments of earnings growth to own firm and competitor innovation (Panel B of Figure 4), these aggregated results do not immediately follow—they also depend on the joint distribution of firm innovation in each industry. Last, we also explore the ability of the model to match how the entire distribution of worker earnings shifts in response to industry innovation. To construct the data equivalents, we re-estimate equation (10), but now replace the dependent variable with percentiles of worker earnings growth calculated in each industry-age-income bin. As we see, the model does quite well in generating an increase in variance and negative skewness of earnings growth in response to industry innovation.

3.3 Model Implications

Our structural model allows us to quantify the welfare cost of the earnings risk borne by workers as well as their willingness to purchase insurance against increases in the rate of industry innovation. To do so, we need to impose some additional structure, namely specify household preferences and asset markets that allow workers to (partly) self-insure against shocks to their labor earnings.

Workers have constant relative risk aversion preferences. They maximize

$$\max_{c} \; \mathrm{E}_0 \int_0^\infty e^{-\rho t} \frac{c_t^{1-\gamma}}{1-\gamma} dt. \tag{20}$$

We choose a value of $\gamma = 5$ as our baseline case. For simplicity, agents are infinitely lived, though results are essentially isomorphic to a case in which agents have a constant probability of death and no utility over bequests, thus ρ can be interpreted as the sum of an agent's rate of time preference parameter and the instantaneous hazard rate of exit. We set $\rho = 4\%$. Workers choose consumption and savings plans subject to the following constraints:

$$d\alpha_{j,t} = (r\,\alpha_{it} + y(s_{j,t}) - c_{j,t})\,dt \tag{21}$$

$$\alpha_{j,t} \geq 0. \tag{22}$$

In particular, workers can save (but not borrow) at an exogenous interest rate r. We choose $r = 2\%$ to match the historical level of the (real) risk-free rate. Exiting households pass their asset holdings to their children, who start life unemployed. Here, $y(s_{j,t})$ denotes the labor earnings that worker j receives given her Markov state $s_{j,t}$, *after* labor income taxes, UI, and other transfers. Following Heathcote, Storesletten, and Violante (2017), we assume that income after tax and transfers is

$$y(s_{it}) = a\, w(s_{j,t})^{1-b}, \tag{23}$$

where $w(s_{j,t})$ captures labor earnings including unemployment benefits. We set $b = 0.181$ following Heathcote et al. (2017), and set a so that the budget is balanced. For simplicity we assume that interest income is subject to a flat income tax, so we can interpret r as the net of tax real interest rate. To reduce the dimensionality of the state space, we assume that UI payments are set at a fraction $\phi = 1/2$ of the lowest possible wage earned by an employed skilled worker. Panel B of Table 2 summarizes our parameter choices.

Given this structure, our problem reduces to a standard incomplete markets consumption-savings problem with post-tax income that follows an exogenous Markov chain. Let $V(\alpha, i)$ denote the value function given current wealth α and that the agent is in income state i. The value function solves the system of Hamilton-Jacobi-Bellman equations for all states $i = 1, \ldots, K$,

$$0 = \max_c \left\{ u(c) - \rho\, V(\alpha, i) + V_\alpha(\alpha, i)\Big(\alpha\, r + y(i) - c\Big) + \sum_{j \neq i} \psi_{i,j}\Big(V(\alpha, j) - V(\alpha, i)\Big) \right\}, \tag{24}$$

subject to

$$c - y(i) \leq 0, \quad \text{if} \qquad \alpha = 0. \tag{25}$$

Here, $\psi_{i,j}$ is the Poisson rate of transitioning from state i to state j. We solve for the value function numerically and then compute the ergodic distribution of assets for each income level by solving the Fokker-Plack equation (see, e.g., Achdou, Han, Lasry, Lions, and Moll, 2020). All details are relegated to Appendix B.2.

We first examine the welfare cost of technological innovation borne by skilled workers. A relatively standard way of expressing welfare cost is in terms of a permanent subsidy/tax on consumption which would make the person indifferent if she were to experience a one standard deviation increase in own firm (λ_f), competitor ($\lambda_{f'}$), or industry innovation (average of λ_f and $\lambda_{f'}$). Suppose the household is currently in state i. Using the fact that the utility function is homogenous of degree $1-\gamma$, if we multiply

consumption by a factor of $1+\Delta$ per period, her new value in state j (post-shock) would be equal to

$$\tilde{V}(\alpha, j; \Delta) = \exp[(1-\gamma)\Delta]V(\alpha, j). \tag{26}$$

Thus, the proportional tax or subsidy Δ which leaves the household indifferent ($\tilde{V}(\alpha, j; \Delta) = V(\alpha, i)$) is equal to

$$\Delta = \frac{1}{1-\gamma} \log\left(\frac{V(\alpha, i)}{V(\alpha, j)}\right). \tag{27}$$

Innovation is associated with significant utility losses for top workers, even in the presence of progressive taxation and ability to self-insure by accumulating assets. The average worker in the model would need to receive a proportional subsidy of $\Delta = 1.5\%$ in her lifetime consumption to offset the utility loss resulting from a one standard deviation increase in industry innovation (average of λ_f and $\lambda_{f'}$). Decomposing this estimate into the welfare cost of own firm (λ_f) and competitor ($\lambda_{f'}$), we find that these utility losses are primarily, though not exclusively, driven by competitor innovation. Specifically, the worker would need to receive a perpetual subsidy of $\Delta = 1.8\%$ to offset her utility losses resulting from a one standard deviation increase in $\lambda_{f'}$. Importantly, even though own-firm innovation λ_f is associated in an increase in average earnings (recall Panel B.i in figure 4), the increase in the left tail is still significant enough to generate a modest utility loss: the required consumption subsidy to offset a one-standard-deviation increase in λ_f is equal to $\Delta = 0.3\%$. In appreciating the magnitude of these estimates, it is important to keep in mind that they correspond to utility losses following a single shock to λ (scaled to unit standard deviation, annualized) rather than shutting down all fluctuations in λ. As such, these estimates imply a substantially higher welfare cost of innovation for (top) workers than the welfare cost of business cycles due to job displacement computed by Krebs (2007).

A natural question to ask in this setting is what would be the willingness of workers to invest in assets that (partially) hedge their earnings risk. A recent literature in asset pricing has argued that 'value' (low Tobin's Q) and 'growth' (high Tobin's Q) firms have differential exposure to technology innovations Papanikolaou (2011); Garleanu et al. (2012); Kogan et al. (2020). For instance, in Papanikolaou (2011) and Kogan et al. (2020), technological progress is partially embodied in new types of capital. As new vintages of capital improve in productivity, growth firms appreciate (as their growth prospects improve) whereas value firms decline (as the price of the capital they own declines). Though our setting here is quite different, one can envision constructing the equivalent of value and growth firms in our model based on their future prospects (λ) and their assets in place (the products they currently produce).

A sufficient statistic that describes a worker's willingness to hedge her earnings risk is the elasticity of her marginal utility to a one-standard-deviation increase in λ.[14] To a first approximation, this elasticity is equal to (minus) the Sharpe ratio (market price of risk) of an investment strategy whose return is maximally correlated with a shock to λ (see, e.g., Cochrane, 2001, for a textbook reference). We can use this idea to compute the Sharpe ratio of such a hypothetical asset if it were introduced, while holding consumption allocations fixed.

Focusing on the average worker, we find that a one standard deviation increase in industry innovation leads to a 0.38 log point increase (on average) in the marginal utility of skilled workers. As before, this is mainly driven by the desire to insure against innovation by competitors (0.426) more than innovation by their own firm (0.057). To put this number into perspective, the model of Kogan et al. (2020) needs to generate approximately a 0.5 log point increase in households' stochastic discount factor (the equivalent of marginal utility in our setting) in order to match the properties of value and growth firms in the data. Kogan et al. (2020) generate this increase in marginal utility assuming an incomplete market for ideas, preferences for relative consumption, and a relatively high degree of risk aversion (57). By contrast, we show that once we allow for displacement of human capital, one can generate similar implications about the price of risk of technology shocks with a much more moderate level of risk aversion (5).

Figure 6 illustrates how these costs vary in the cross-section of workers. Panels A and B illustrate that there is significant heterogeneity among the workers in the model in the degree through which they are affected. The next set of panels illustrate that a significant part of the heterogeneity is determined by the worker's current income level. In particular, as we see in Panels C and D of Figure 6, these costs are significantly larger for more highly paid workers. Workers that are more highly paid in the model have the most to lose, and hence they are willing to pay more to insure against these risks. By contrast, Panels E and F illustrate that there is no significant gradient across asset holdings. Since asset holdings are endogenous, workers that are most affected aim to accumulate a higher level of liquid assets in order to self-insure against these shocks.

In sum, even though the model refers to skilled (top) workers as a group, the model still generates heterogenous responses as a function of income within that group. To illustrate this point, Figure 7 revisits the model response of workers' income risk to own firm and competitor innovation shocks

[14] In the model equilibrium computed at the estimated parameters, the stationary distribution of assets does not include any probability mass near the borrowing constraint. As such, the marginal utility of consumption equals the marginal value of wealth essentially with probability one. Hence, we use these terms interchangeably in our discussions.

(Panel B of Figure 4) but now conditions on the worker's (pre-shock) level of income. Examining the figure, we note the higher sensitivity of the left tail of earnings growth (to either innovation shock) for the workers that are above versus below the median income. This pattern is qualitatively similar to Panels C and D of Figure 3.

4 Additional Model Implications

The model in Section 3 generates a relation between the degree of technological innovation and labor income risk. Depending on whether the innovations originate in the same firm as the worker is employed or in other firms in the industry, the model generates shifts in the distribution of earnings growth for top workers that is consistent with the data.

Here, we dig deeper into the model mechanism along two dimensions. First, the increase in the left-tail of earnings growth in the model primarily operates through worker separations. Section 4.1 explores the extent to which this model prediction aligns with the data. Second, the model emphasizes innovations that are likely to displace worker skills; this mechanism generates increases in the left tail of workers earnings growth for top workers. Section 4.3 accordingly focuses on measures of firm innovation that are more likely to be novel or represent process improvements.

4.1 Worker Mobility and Earnings Risk

A key mechanism through which the model generates increases in the left tail of earnings growth is worker separations. To illustrate this point, we next decompose the model results in Panel B of Figure 4 into stayers and movers; movers are defined as workers who leave the firm within the first year following the shock. We then trace out the implied marginal effect of a one standard deviation change in λ_f separately for stayers and movers. We plot the results in Figure 8 (red bars).

Examining the red bars of Figure 8 we see that the increased risk in worker earnings, in response to either innovation by their own firm or its competitors, is primarily borne by workers who leave the firm. In the model, worker separations are costly because, while unemployed, workers miss out on earnings increases, and when re-employed, start at the lowest level of worker productivity $\bar{z}$. That is, workers that leave the firm are falling off the job ladder. Innovation affects the likelihood of job loss through two key mechanisms. First, own firm innovations can displace the worker's human capital—workers draw a new productivity level according to (18). Workers whose productivity falls below $\bar{z}$ are replaced.

Second, innovation by the firm's competitors can result in the loss of the firms' position as a leading producer and hence to worker separations. An increase in the rate of innovation by the firm, or its competitor, λ increases the likelihood of these events occurring and therefore expands the set of workers potentially affected.

We next examine whether these patterns also obtain in the data. To do so, we estimate equation (12) separately for workers that move (movers) versus those that do not (stayers). As before, we focus on cumulative growth rates between t and $t+5$ using (2). In contrast to the model, separations in the data are likely not instantaneous. As such, we want to allow for delayed effects on mobility, but at the same time capture worker earnings changes in the new job. We define movers as those workers who 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 $t+3$. We plot the results in the data in Figure 8 (blue bars) for the top workers; Figure 9 presents our findings across all income groups.

Comparing the red (model) and blue (data) bars in Figure 8, we note that the model's implications regarding the link between innovation and the left tail of earnings growth of top workers are consistent with the data. In particular, we note 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. Similarly, we see 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, in both the model and the data, the magnitude of the increase in the left tail is considerably larger for exiting workers. Examining Figure 9, we note that even though the increase in the left tail among exiting workers is present across all income groups, it is significantly larger in magnitude for top workers.[15]

Last, focusing on the differences in the right tail of earnings growth between movers and stayers in Figure 8, we note that the model generates a somewhat larger increase in the right tail for movers than for

[15] An alternative interpretation of these results is that they reflect adverse selection (Gibbons and Katz, 1991). 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. 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.10 shows that workers that left the firm following periods of high firm/industry stock returns do not experience more negatively-skewed income growth than workers that left the firm during periods of low stock returns. For adverse selection to explain the contrast between Figure 3 and Figure 11, it has to be the case that innovation is somehow different than other shocks which affect the firm. Our model mechanism can be interpreted in this light: 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 the lack of skills that have become relevant.

stayers relative to the data. Conditional on staying with the firm, an increased rate of new innovations by potential entrants lowers incumbent firms' profits (markups), which has an unambiguously negative impact on incumbent workers' earnings. Hence, the whole distribution of earnings growth rates shifts to the left for stayers. By way of contrast, some workers who switch jobs after being displaced by competitors' innovations end up being matched with more productive, better paying firms in unrelated product lines with positive probability. This force leads to an expansion of the right tail for movers.

4.2 Innovation and unemployment risk

Our model links innovation to workers earnings risk through the (potential) displacement of workers' human capital. In both the data and the model, exiting workers experience a substantially more negatively-skewed distribution of earnings growth in response to innovation outcomes than continuing workers. In the model, this pattern occurs through an increased likelihood of job separation; the length of an unemployment spell is largely independent of the worker's characteristics. In the data, however, this pattern can be driven by both an increased likelihood of separation and an increased length of the unemployment spell. Since both of these factors will determine the increase in the left tail in the data, we next examine them jointly.

We construct a measure of long-term unemployment based on the number of years with zero W-2 earnings.[16] 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. Appendix 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.[17]

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

$$U_{i,t:t+5} = a_h \, A_{f,t} + b_h \, A_{I \setminus f,t} + c_h \, Z_{i,t} + \varepsilon_{i,t}. \tag{28}$$

The vector of controls Z contains the same worker- and firm-level controls as equation (12).

Panel A of Table 3 presents the results. We see a modest but statistically significant link between innovation and unemployment risk. As before, magnitudes are larger for top workers. A one standard

[16] Measuring directly the length of unemployment spells is not possible in our data, since we do not observe any information on unemployment benefits.

[17] 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.

deviation increase in $A_{f,t}$ is associated with a 0.011 increase in number of years with zero W-2 earnings (a sizable effect given a mean of 0.142). Similarly, a one standard deviation increase in competitor innovation is associated with up to a 0.022 increase in the number of years without employment for the firm's top workers. [18]

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) discuss the secular increase in the number of these 'conditional applicants' since 1984, partly as a response to changes in determination standards, but also more importantly, as a response to changes in labor market conditions. In the context of our setting, workers may choose to exit the labor market if their skills have depreciated sufficiently. Further, risk may play an important role: the decision to accept transfer payments involves exchanging a claim on a stream of (comparatively) risky future labor earnings for a known, safe stream of transfers. An increase in the riskiness of labor earnings decreases the attractiveness of the labor market.

We therefore re-estimate (28), but now the dependent variable $U_{i,t:t+5}$ takes the value of one if the worker i has filed for disability insurance sometime between year $t+1$ and $t+5$. Panel B of Table 3 shows that there is a small but statistically significant link between innovation and the likelihood of applying for DI benefits. As before, these effects are stronger in magnitude for skilled workers. A one standard deviation increase in the rate of innovation by the firm (its competitors) is associated with a 0.08 (0.19) percentage point increase in the likelihood of applying for DI. To evaluate these magnitudes, it is useful to compare to the unconditional mean of the dependent variable. As we see in Panel A of Table A.1, applying for disability insurance is not an altogether rare phenomenon. Approximately 2.6% of workers apply for disability insurance over a 5-year period.

[18] These results shed some light on the patterns in Figure 9. Specifically, they suggest that part of the large increase in the left tail for movers following innovation outcomes we document in Figure 9 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.14 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, there is no increase in the left tail of earnings growth for top workers in response to innovation by their own firm. The increase in left tail in response to competitor innovation is still present, though significantly smaller in magnitude.

4.3 Heterogenous Innovations

In the model developed in Section 3, innovation consists of new production methods or procedures that help the firm lower production costs. A key part of the model mechanism is that incumbent worker's productivity may be specific to a technology vintage. As such, the model's implications regarding the response of worker earnings to own firm innovation are more likely to apply to specific types of improvements: innovations that are more novel, or distinct, relative to what the firm has done in the past; and, innovations that represent process, rather than product, improvements. To test these model predictions more directly, we next decompose the own-firm innovation measure A_f into these different types of improvements. Since the implications of the model regarding how workers' earnings respond to innovation by competing firms could apply to any type of innovations—they are driven by changes in market share—we focus on the response to own innovation.

4.3.1 Novel vs less novel innovations

We begin by identifying novel patents, that is patents that are distinct from the firm's prior innovations. To do so, we need a distance measure between patents. We use the text-based methodology of Kelly et al. (2020), which we briefly describe next.

A key consideration in constructing a similarity metric for a pair of text documents is to appropriately weigh words by their importance. It is more informative if terms such as 'electricity' and 'petroleum' enter more prominently into the similarity calculation than common words like 'process' or 'inventor.' Further, as technology evolves, some words may become more common (for instance, 'electricity' in the 1900s vs 2000s) and hence it is important to have a dynamic characterization of novelty.

Thus, for a patent pair (i, j) we construct a vector of weights

$$TFBIDF_{w,i,t} = TF_{w,i} \times BIDF_{w,t}, \quad t \equiv \min(i, j) \tag{29}$$

which weighs words according to their relative importance in the patent document, as captured by the first term,

$$TF_{pw} \equiv \frac{c_{pw}}{\sum_k c_{pk}}, \tag{30}$$

as well as their frequency in the corpus of all prior documents, which is captured by the second term,

$$BIDF_{wp} = \log\left(\frac{\#\text{ patents prior to } p}{1 + \#\text{ documents prior to } p \text{ that include term } w}\right). \tag{31}$$

This last component (31) evolves over time as a term becomes more or less widely used, reflecting the history of invention up to, but not beyond, the new patent's arrival.

Next, we compute the pairwise cosine similarity between each patent i and all the patents $j \in J$ that are filed by the same firm over the last five years. We measure the distance between two patents i and j using cosine similarity,

$$\rho_{i,j} = \frac{TFBIDF_{i,t}}{||TFBIDF_{i,t}||} \cdot \frac{TFBIDF_{j,t}}{||TFBIDF_{j,t}||}. \tag{32}$$

Here, $TFBIDF_{i,t}$ and $TFBIDF_{j,t}$ are vectors of length W, which is the size of the set union for words in pair (i, j). Last, we classify a patent i as novel $N_i = 1$, if it is sufficiently distinct to the firm's previous patents,

$$N_i = 1 \Leftrightarrow \max_{j \in P(f)} \rho_{i,j} \leq 0.5. \tag{33}$$

Here, $P(f)$ refers to the set of patents filed by the same firm over the last 5 years. Approximately 35% of all patents are classified as novel under this criterion.

Given our patent-level measure of novelty, we compute

$$A_{f,t}^{novel} = \frac{\sum_{j \in P_{f,t}} \xi_j \, N_j}{B_{ft}}. \tag{34}$$

We term the residual level of innovation as not novel, $A_{f,t}^{not-novel} = A_{f,t} - A_{f,t}^{novel}$. The correlation between A_f^{novel} and $A_f^{not-novel}$ is approximately 22%.

We then estimate a modified version of equation (12) as

$$g_{i,t:t+h} = a_0 + a_\tau^n \, A_{f,t}^{\text{novel}} + a_\tau^o \, A_{f,t}^{\text{not-novel}} + b_\tau \, A_{I \setminus f,t} + c \, Z_{i,t} + \varepsilon_{i,t}. \tag{35}$$

For ease of comparison, marginal effects are scaled by the cross-sectional standard deviation of own firm innovation.

Panels A and B of Figure 10 show that, consistent with the spirit of our model, novel innovations are associated with significantly higher earnings risk for the firm's top workers. Comparing Panels A and B, there are two points worth noting. First, the magnitudes of the estimated worker earnings responses are significantly larger for novel, rather than less novel, innovations. Second, the increase in the left tail of earnings growth is only present in response to novel innovations.

4.3.2 Process vs non-process

We next distinguish between process and non-process innovations. To do so, we use the data and classification procedure of Bena and Simintzi (2019), who identify the fraction θ_j of claims of patent j that can be identified with a process using textual analysis.[19] The mean value of θ_j in our sample is 32%, but there is considerable dispersion as its standard deviation is approximately 38%. The residual claims $1-\theta_j$ can refer to other 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,

$$A_{f,t}^{\text{process}} = \frac{\sum_{j \in P_{f,t}} \xi_j \, \theta_j}{B_{ft}}. \tag{36}$$

We term the residual level of innovation as other, $A_{f,t}^{other} = A_{f,t} - A_{f,t}^{process}$. The correlation between A_f^{proc} and A_f^{other} is approximately 60%.

We next estimate a modified version of equation (12),

$$g_{i,t:t+h} = a_0 + a_\tau^p \, A_{f,t}^{\text{process}} + a_\tau^o \, A_{f,t}^{\text{other}} + b_\tau \, A_{I \setminus f,t} + c \, Z_{i,t} + \varepsilon_{i,t}. \tag{37}$$

As before, we scale the marginal effects by the cross-sectional standard deviation of own firm innovation.

Panels C and D of Figure 10 show that our findings are broadly consistent with the model mechanism. Comparing Panels C (process) and D (non-process), we see that the two types of innovation have a qualitatively different effect on the distribution of earnings growth for the firm's own workers. 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, an increase in $A_{f,t}^{proc}$ is associated with a 4.5 percentage point decrease in the 5^{th} percentile of income growth. These findings are qualitatively consistent with the model predictions—Panel B.i in Figure 4.

Naturally, not all process innovations may be truly novel. Our model would imply that novel process improvements are more likely to be associated with the displacement of top workers. As such, we further decompose process innovations into novel vs not novel,

$$A_{f,t}^{\text{novel–process}} = \frac{\sum_{j \in P_{f,t}} \xi_j \, \theta_j \, N_j}{B_{ft}}, \tag{38}$$

[19] Bena and Simintzi (2019) identify patent claims that refer to process innovation 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.

where $N_j = 1$ defines a novel innovation following the discussion immediately above. Accordingly, we re-estimate (37) above, now decomposing $A_{f,t}^{proc}$ into novel $A_{f,t}^{\text{novel–process}}$ and non-novel $A_{f,t}^{proc} - A_{f,t}^{\text{novel–process}}$ process innovation. Panels E and F of Figure 10 show that, consistent with our model, novel process innovations, in particular, are more likely to be associated with an increase of the left tail of earnings growth for the firm's top workers.

Last, we revisit the analysis in Section 4.1 using this distinction between process and non-process improvements. In Appendix Figure A.6, we show that, consistent with our model, these left tail effects for process innovation are considerably larger for workers who subsequently leave the firm. By contrast, estimates for non-process innovations are fairly similar between movers and stayers. Further, Appendix Figure A.22 repeats this exercise allowing for coefficients to vary with worker tenure as well as mobility. Though worker tenure has no independent role in the model, we may expect that workers that have worked longer at the firm are more likely to have acquired specific skills and are therefore more at risk of displacement. Consistent with this view, we find that these left tail effects for top workers are most pronounced for those with 3 or more years of tenure with the firm, especially those who leave.

5 Additional Results and Robustness Checks

In this section, we discuss several additional results which speak to the robustness of our findings.

As noted above, the results in Figure 2 suggest that our innovation measures do not pick up underlying, pre-existing trends in firm-specific profitability. However, innovation is partly related to firms' decisions to invest in R&D. Unobservable characteristics may jointly drive firms' willingness to invest in R&D, profits, and the distribution of worker earnings. 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 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.8 in the Appendix shows that we find that controlling for R&D spending does not materially affect the magnitude of the estimated coefficients.

The results in Figure 3 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. 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 view is 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 3).

Third, the fattening of the left tail that we document in Figure 3 appears to be specific to innovation outcomes. Specifically, 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. Second, we repeat this exercise where we instead condition on realized profit growth over the next five years, both by the firm and its competitors. In the case of the firm, this measure given by equation (8), and we aggregate competitor innovation analogously to (4). Figure 11 presents the results.

Contrasting Figure 11 to Figure 3, we see that the relation of the earnings of top workers with innovation outcomes is qualitatively distinct from the relation with other sources of firm profitability. In Panel A, we see that a positive shock to the firm's stock price is associated with a mostly 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 3. 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

[20] Appendix Figure A.10 further confirms our conclusion: we find no analogous fattening of the left tail for workers who move following periods with high firm stock returns.

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. Appendix Tables A.4 and A.5 illustrate that these qualitative differences also arise when we focus on conditional mean (OLS) results.

Last, another possibility is that the higher left-tail sensitivity for higher income workers is really picking up a worker age effect rather than a highly-paid worker effect. In particular, even though worker earnings are defined net of life cycle dummies—see equation (1)—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 examine if this is the case, we estimate equation (12) separately for workers of different age groups and income levels. Indeed, Appendix Figure A.16 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 Conclusion

Our analysis reveals several new facts regarding how the distribution of worker earnings growth changes following technological improvements by firms (and their 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. Our findings imply that 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.

[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.20 to A.23, continue to hold within tenure groups, and displacement effects are somewhat more pronounced for higher tenure workers (especially those that move).

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 (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy 128*(6), 2188–2244.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2020). Income and wealth distribution in macroeconomics: A continuous-time approach. *Review of Economic Studies forthcoming*.

Adão, R., M. Beraja, and N. Pandalai-Nayar (2020). Technological transitions with skill heterogeneity across generations. Technical report, National Bureau of Economic Research.

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

Aghion, P., U. Akcigit, A. Deaton, and A. Roulet (2016). Creative destruction and subjective well-being. *American Economic Review 106*(12), 3869–97.

Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen (2018). On the returns to invention within firms: Evidence from Finland. In *AEA Papers and Proceedings*, Volume 108, pp. 208–12.

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.

Akcigit, U., J. Grigsby, and T. Nicholas (2017). The rise of American ingenuity: Innovation and inventors of the Golden Age. Technical report, National Bureau of Economic Research.

Atalay, E., P. Phongthiengtham, S. Sotelo, and D. Tannenbaum (2020, April). The evolution of work in the United States. *American Economic Journal: Applied Economics 12*(2), 1–34.

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

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019). Do tax cuts produce more Einsteins? the impacts of financial incentives versus exposure to innovation on the supply of inventors. *Journal of the European Economic Association*.

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

Braxton, J. C. and B. Taska (2020). Technological change and the consequences of job loss. Working paper, University of Minneapolis.

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.

Cochrane, J. H. (2001). *Asset Pricing*. Princeton University Press.

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.

Deming, D. J. and K. Noray (2020, 06). Earnings Dynamics, Changing Job Skills, and STEM Careers*. *The Quarterly Journal of Economics 135*(4), 1965–2005.

Eisfeldt, A. L., E. Kim, and D. Papanikolaou (2020, November). Intangible value. Working Paper 28056, National Bureau of Economic Research.

Ellul, A., M. Pagano, and F. Schivardi (2017, 09). Employment and Wage Insurance within Firms: Worldwide Evidence. *The Review of Financial Studies 31*(4), 1298–1340.

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.

Frydman, C. and D. Papanikolaou (2018). In search of ideas: Technological innovation and executive pay inequality. *Journal of Financial Economics 130*(1), 1–24.

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.

Grossman, G. M. and E. Helpman (1991). Quality ladders in the theory of growth. *Review of Economic Studies 58*(1), 43–61.

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.

Heathcote, J., K. Storesletten, and G. L. Violante (2017, 06). Optimal Tax Progressivity: An Analytical Framework*. *The Quarterly Journal of Economics 132*(4), 1693–1754.

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.

Huckfeldt, C. (2018). Understanding the Scarring Effect of Recessions. Technical report.

Ingram, B. F. and B.-S. Lee (1991, February). Simulation estimation of time-series models. *Journal of Econometrics 47*(2-3), 197–205.

Jarosch, G. (2021, February). Searching for job security and the consequences of job loss. Working Paper 28481, National Bureau of Economic Research.

Jones, C. I. and J. Kim (2018). A schumpeterian model of top income inequality. *Journal of Political Economy 126*(5), 1785–1826.

Jovanovic, B. and Y. Nyarko (1996). Learning by doing and the choice of technology. *Econometrica 64*(6), 1299–1310.

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

Klette, T. J. and S. Kortum (2004). Innovating inds and aggregate innovation. *Journal of Political Economy 112*(5), 986–1018.

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).

Krusell, P., L. E. Ohanian, J.-V. Ríos-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: A macroeconomic analysis. *Econometrica 68*(5), 1029–1053.

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 U.S. labor market. *Working Paper*.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics 13*(4), 653–677.

Papanikolaou, D. (2011). Investment shocks and asset prices. *Journal of Political Economy 119*(4), 639–685.

Park, H. (2019). An intangible-adjusted book-to-market ratio still predicts stock returns. *Critical Finance Review 25*(1), 207–236.

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

Poterba, J. M. and A. A. Samwick (1995). Stock Ownership Patterns, Stock Market Fluctuations, and Consumption. *Brookings Papers on Economic Activity 26*(2), 295–372.

Sampat, B. and H. L. Williams (2019, January). How do patents affect follow-on innovation? Evidence from the human genome. *American Economic Review 109*(1), 203–36.

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.

Toivanen, O. and L. Väänänen (2012). Returns to inventors. *Review of Economics and Statistics 94*(4), 1173–1190.

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.

Table 1: Innovation, firm profitability and worker earnings

	A. Firm Profitability	B. Worker Earnings					
		All Workers	Earnings Rank				
			Bottom [0,25]	[25,50]	[50,75]	[75,95]	Top [95,100]
Firm Innovation, market value (A_f^{sm})	7.99	1.38	1.19	1.09	1.40	1.85	1.53
	(7.39)	(15.46)	(10.86)	(8.07)	(9.98)	(14.41)	(8.34)
Implied Elasticity		0.173	0.149	0.137	0.176	0.231	0.191
Competitor Innovation, market value ($A_{I\setminus f}^{sm}$)	-4.93	-1.88	-1.92	-1.40	-1.01	-2.20	-5.92
	(-7.81)	(-5.42)	(-6.41)	(-5.68)	(-3.91)	(-6.87)	(-11.90)
Implied Elasticity		0.381	0.39	0.284	0.204	0.444	1.201

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 (3) 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 (4)). 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. In addition to results for all workers (homogenous coefficients), in the last five columns we allow for heterogenous coefficients for workers with different earnings levels and estimates are scaled to correspond with a one standard deviation change in each variable of interest; t-statistics are in parentheses. The worker earnings rank is defined net of deterministic life-cycle effects.

Table 2: Model Parameters

Description	Parameter	Value
A. Parameters governing earnings process		
Drift of (log) worker efficiency if innovation occurs	μ_z	0.013
Volatility of (log) worker efficiency if innovation occurs	σ_z	0.009
Persistence of (log) worker efficiency if innovation occurs	ρ_z	0.413
Innovation rate (high state)	λ_H	1.772
Innovation rate (low state)	λ_L	0.299
Innovation step size ($\times$ 100)	h	0.060
Transition rate into high innovation state	μ_H	0.129
Transition rate into low innovation state	μ_L	0.082
Rate of radical innovation by competing firm	p	0.056
Volatility of initial technology gap when incumbent overtakes	σ_k	0.007
Job finding rate	θ	3.091
B. Household preferences, taxes and transfers, and asset markets		
Rate of time preference (including death hazard rate)	ρ	0.050
Risk aversion	γ	5
Tax progressivity	b	0.181
UI benefits as fraction of lowest wage (among managers)	ϕ	0.500
Interest rate	r	0.020

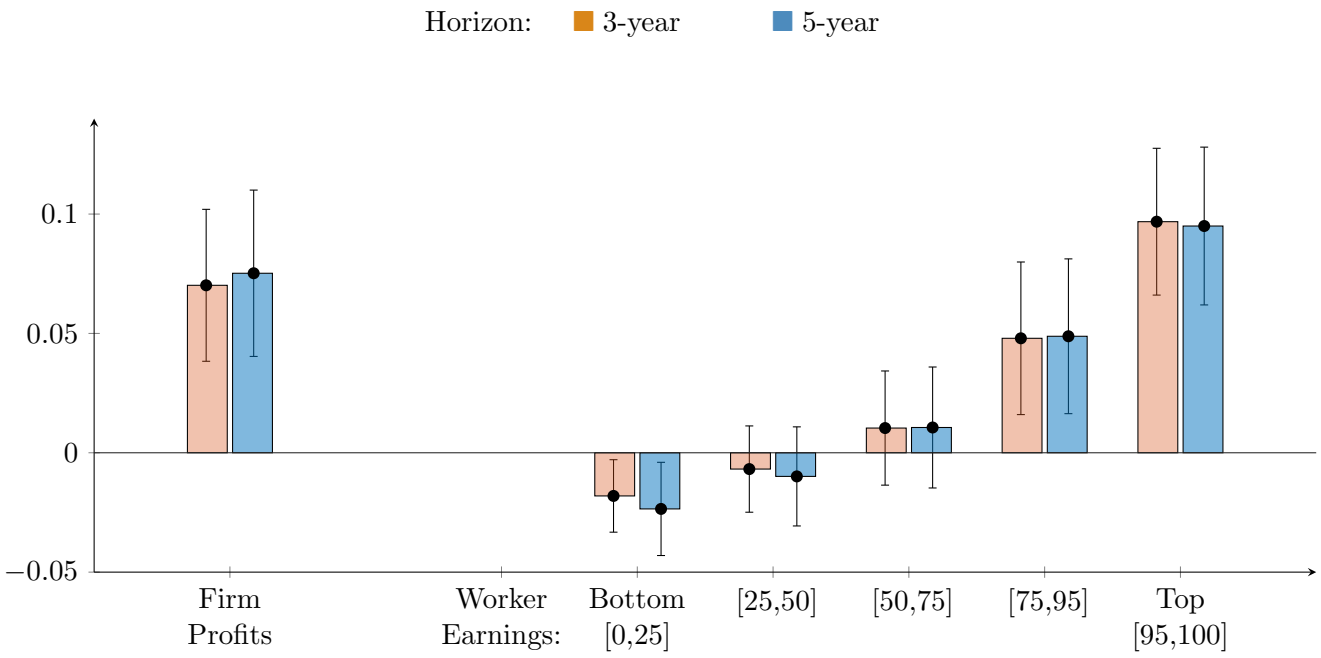
Note: Table reports the parameter estimates used in our model solution. The parameter estimates in Panel A are obtained by minimizing equation (19) in the main text, which computes the distance in targeted statistics between the model and the data. Figure 4 reports the targeted moments along with model fit. See the main text for a discussion of our parameter choices in Panel B.

Table 3: Innovation and long-term unemployment

A. Number of years unemployed, 5yr horizon					
Innovation	Worker earnings rank				
	[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	-0.28	0.55	0.43	0.61	1.06
	[-2.58]	[6.32]	[4.96]	[4.40]	[5.44]
Innovation by competitors, $A_{I\setminus f}$	-0.73	0.49	1.42	2.07	2.17
	[-3.55]	[3.08]	[9.20]	[10.33]	[8.36]
B. Application for disability insurance (DI), 5yr horizon					
Innovation	Worker earnings rank				
	[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	-0.11	0.02	0.04	0.05	0.08
	[-6.56]	[1.46]	[3.92]	[6.03]	[7.05]
Innovation by competitors, $A_{I\setminus f}$	-0.35	-0.16	0.02	0.18	0.19
	[-8.70]	[-5.41]	[0.76]	[7.35]	[5.87]

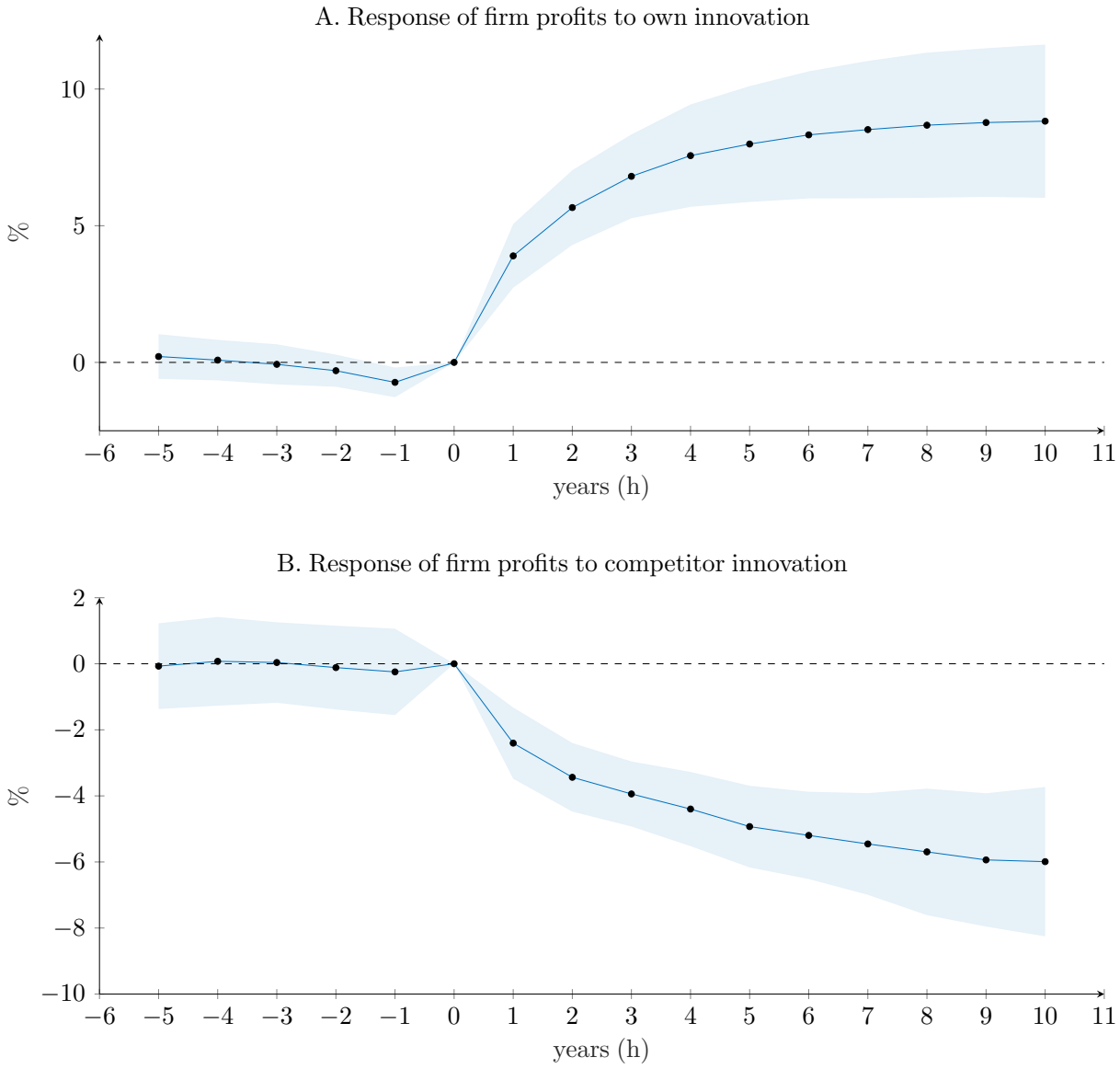
Note: Table reports the relation between innovation and long-term unemployment. Specifically, we report the key coefficient estimates of equation (28) in the paper. In Panel A, the dependent variable is a count of the number of years of zero W2 earnings worker i has experienced between years $t+1$ and $t+5$, which is then multiplied by 100. In Panel B, 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 $t+5$ ($\times 100$). 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 are clustered at the firm level at t-statistics are displayed in parentheses.

Figure 1: Industry Innovation and Risk in Firm Profits and Worker Earnings



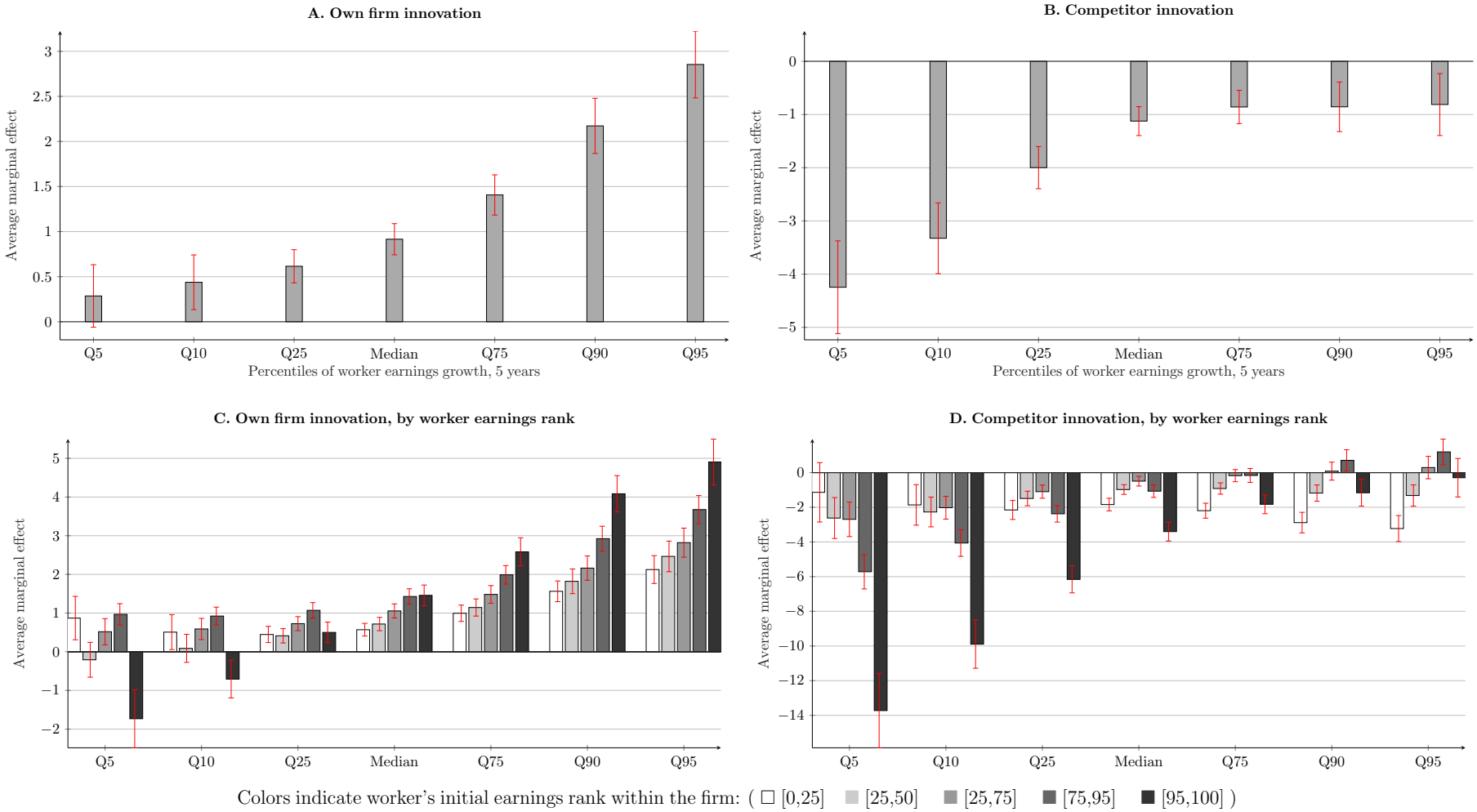
Note: This figure shows the relation between industry innovation and risk, as measured by the log cross-sectional dispersion (variance) of firm profits and worker earnings growth. In particular, we plot the estimated β coefficients from equations (9) and (10) in the main text, which capture responses of the log variance of the conditional distribution of firm and worker earnings growth, respectively, to a one standard deviation industry innovation. In addition, we report separate estimates for different groups of workers who are ranked according to their percentiles of the distribution within each industry and age bin. Worker earnings are defined net of deterministic life cycle effects based on equation (1) in the main text.

Figure 2: Firm profits and innovation across horizons



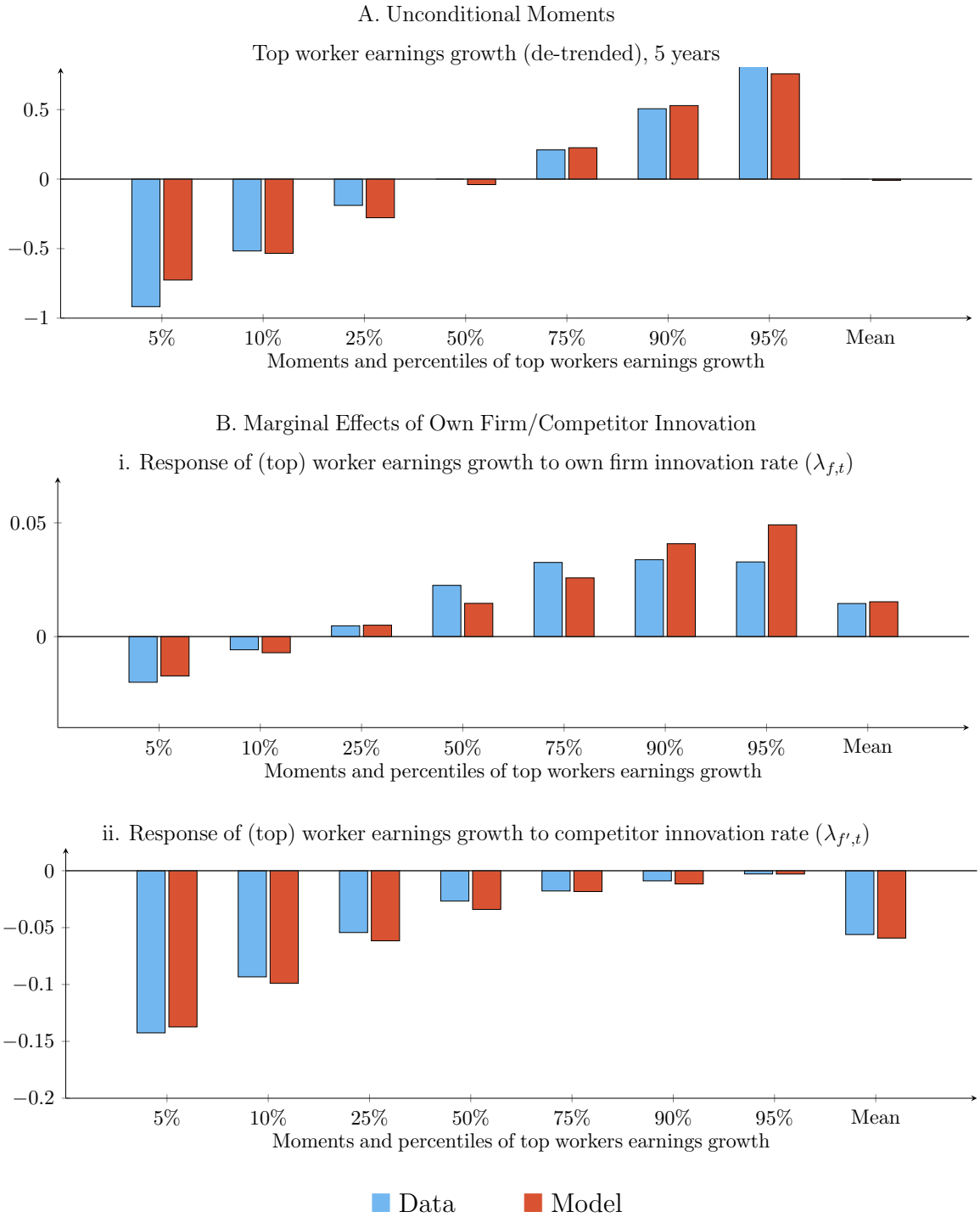
Note: This figure shows the relation between firm profits and own/competitor innovation across different horizons. In particular, we plot the estimated coefficients a_h (Panel A) and b_h (Panel B) from equation (11) in the main text, as we vary the horizon over which compute profits from $h = -5$ to $h = 10$. The estimated coefficient a_h (b_h) corresponds to the relation between the firm's average profits between years $t + 1$ to $t + h$ and and a one standard deviation change in own firm (competitor) innovation at time t; negative values for h correspond to average profits prior to time t, e.g. from $t - h$ to $t - 1$. Bars correspond to 95% confidence intervals computed with standard errors that are clustered by firm and time.

Figure 3: Innovation and Worker Earnings Risk



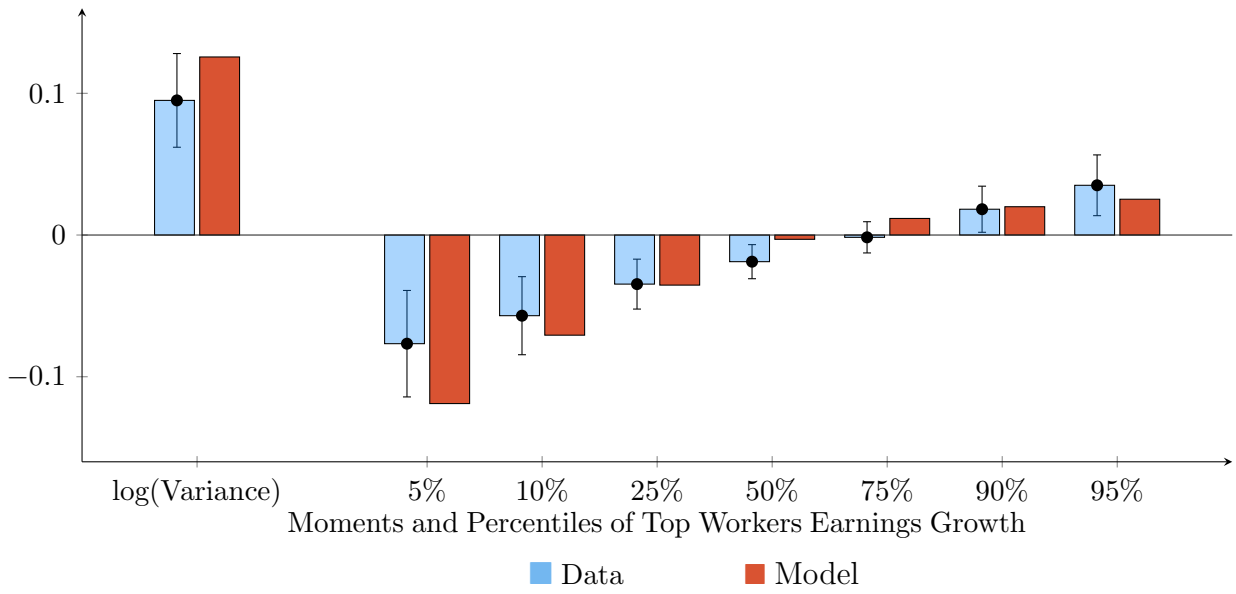
Note: This figure compares the average marginal effects (implied by the quantile regression, equation (12) in the main text) of own firm (left column) and competitor (right column) innovation on quantiles of income growth. The bottom row allows the coefficients to vary by different levels of initial worker earnings (shaded). We focus on 5-year worker earnings growth rates. Estimates are scaled to correspond with a one standard deviation change in each variable of interest. The worker earnings rank is defined net of deterministic life cycle effects. The units on the vertical axis correspond to log points (times 100).

Figure 4: Estimation Targets and Model Fit



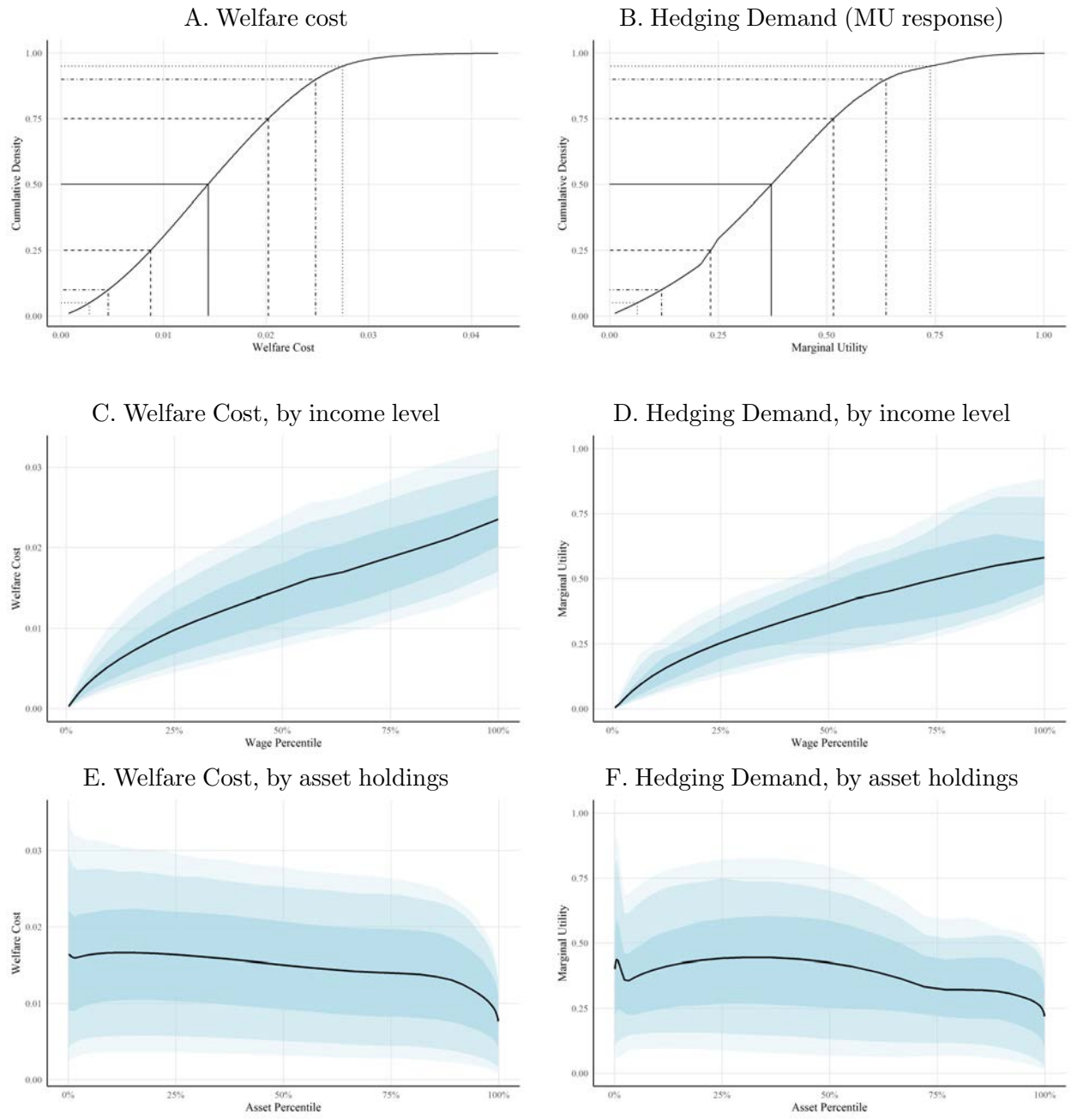
Note: This figure reports the list of targets used to estimate the model parameters (in Panel A of Table 2) and summarizes the model fit to the data. Panel A includes unconditional moments of (top) worker earnings growth. Panel B compares the average marginal effects of firm—and competitor—innovation in simulated data from the model (red) versus the data (blue). The data analogue correspond to the values for top workers (top 95%) in Figure 3. We compute responses of quantiles of earnings growth to a one-unit standard deviation increase in λ_f (own innovation) and $\lambda_{f'}$ (competitor innovation); given that λ is a two-point process, we examine the response from λ_L to λ_H, scaled by the standard deviation λ. Prior to the shock, workers are drawn from the ergodic distribution. We focus on 5-year growth rates. The units on the vertical axis correspond to log points.

Figure 5: Industry Innovation and Individual Earnings Risk: Model vs Data



Note: This figure compares the relation between innovation and income risk in the model versus the data. Specifically, we compare the estimated response of log variance and percentiles of income growth to a one standard deviation change in average industry innovation in simulated data from the model (red) versus the data (blue) at a five year frequency. The data analogues correspond to the values for top workers (top 95%) in Figure 1. Data analogues for conditional quantiles are estimated by estimating (10) using conditional percentiles, rather than log variances, as the dependent variable for workers within the top 5% of each industry-age-year cell. In the model, average industry innovation is $(\frac{1}{2}\lambda_f + \frac{1}{2}\lambda_{f'})$; given that λ_f and $\lambda_{f'}$ follow two-point processes, we compute the response associated with both λ_f and λ_f' increasing from λ_L to λ_H, scaled by the standard deviation of average λ. See notes to Figure 4 for additional details. Table 2 summarizes our parameter choice.

Figure 6: Impact of Industry Innovation Risk on Worker Welfare

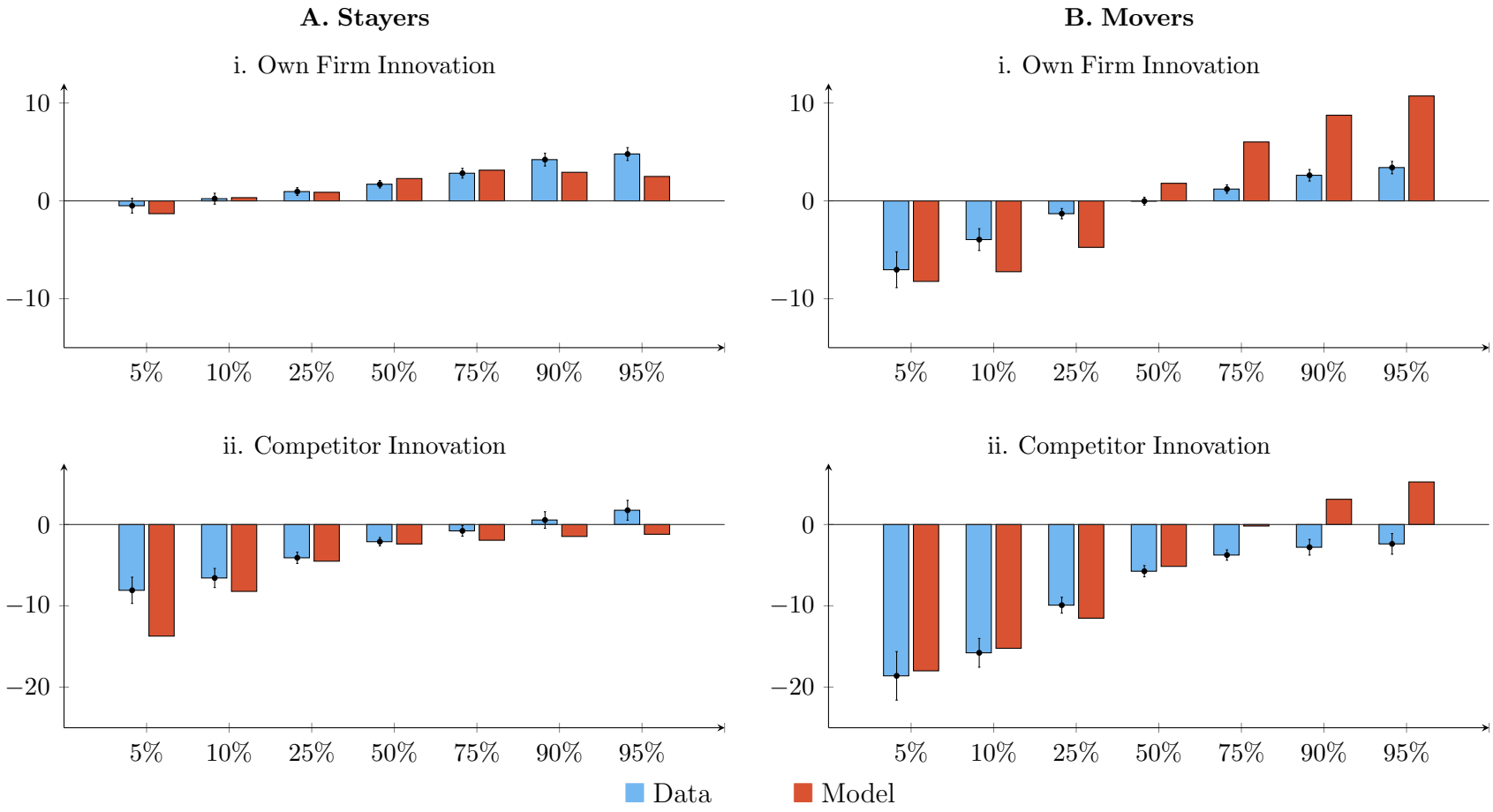


Note: This figure summarizes the welfare cost of innovation in the model, along with agents' willingness to hedge this risk, as summarized by increases in their marginal utility. Our primary focus is on an increase in industry innovation ($\frac{1}{2}\lambda_f + \frac{1}{2}\lambda_{f'}$). The left panel reports the certainty equivalent utility loss (measured in units of per-period consumption) and in the right panel the hedging demand (the expected log change in marginal utility following the shock) associated with this event, scaled to correspond with a 1 standard change in average industry innovation. Given that λ_f and $\lambda_{f'}$ follow two-point processes, we compute responses associated with both λ_f and λ'_f increasing from λ_L to λ_H, then rescale by the standard deviation of average λ. In the top row, we report the CDF of the cross-sectional distribution of these welfare measures, drawn from the ergodic distribution of the model for workers with $\lambda_f = \lambda'_f = \lambda_L$. In the next two rows, we summarize the median (solid lines) and conditional 5th, 10th, 25th, 75th, 90th, and 95th percentiles (shaded bands) of the conditional distribution of these measures as a function of assets and income, respectively. See notes to Figure 4 for additional details. Table 2 summarizes our parameter choice.

Figure 7: Model: Innovation Risk Exposure, by (skilled) worker earnings level

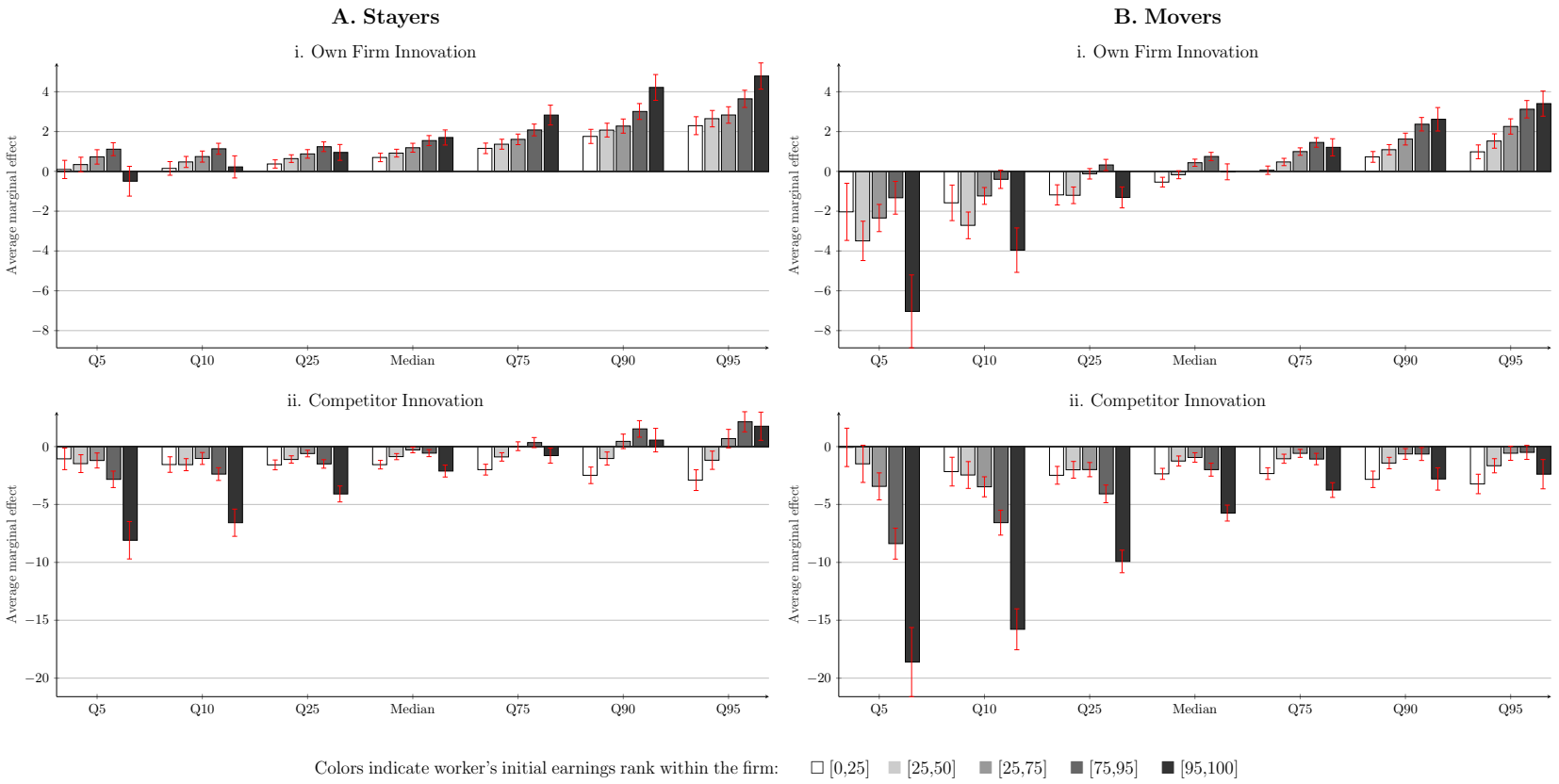
Note: This figure illustrates the extent to which the exposure to innovation varies with the earnings level of (skilled) workers in the model. Specifically, we plot the average marginal effects of firm (Panel A) and competitor (Panel B) innovation in simulated data from the model. We compute responses of quantiles of earnings growth to a one-unit standard deviation increase in λ_f (own innovation) and $\lambda_{f'}$ (competitor innovation); given that λ is a two-point process, we examine the response from λ_L to λ_H, scaled by the standard deviation λ. We examine outcomes separately for workers whose (pre-shock) wage earnings are above the median (red) versus below the median (orange). Prior to the shock, workers are drawn from the ergodic distribution. We focus on 5-year growth rates. The units on the vertical axis correspond to log points (times 100). See notes to Figure 4 for additional details. Table 2 summarizes our parameter choice.

Figure 8: Earnings Risk and Innovation: Movers versus Continuing Workers



Note: This figure compares the innovation earnings exposure faced by continuing (Panel A) and moving (Panel B) workers in the model and the data. Red bars correspond to simulated data from the model. Blue bars correspond to the average marginal effects of firm/competitor innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers at the top of the earnings distribution (top 95% in terms of income relative to other workers in the same firm). We examine outcomes separately for workers that remain with the firm (stayers) versus workers that leave the firm (movers). In the data, movers/stayers are defined depending on whether they remain with the same firm within the next three years; see discussion in the main text for more details. In the model, movers are defined as workers who leave the firm the year following a change in λ_f or $\lambda_{f'}$. We focus on 5-year growth rates. The units on the vertical axis correspond to log points (times 100). See notes to Figure 4 for additional details. Table 2 summarizes our parameter choice.

Figure 9: Earnings Risk and Innovation: Movers versus Continuing Workers (all worker levels)

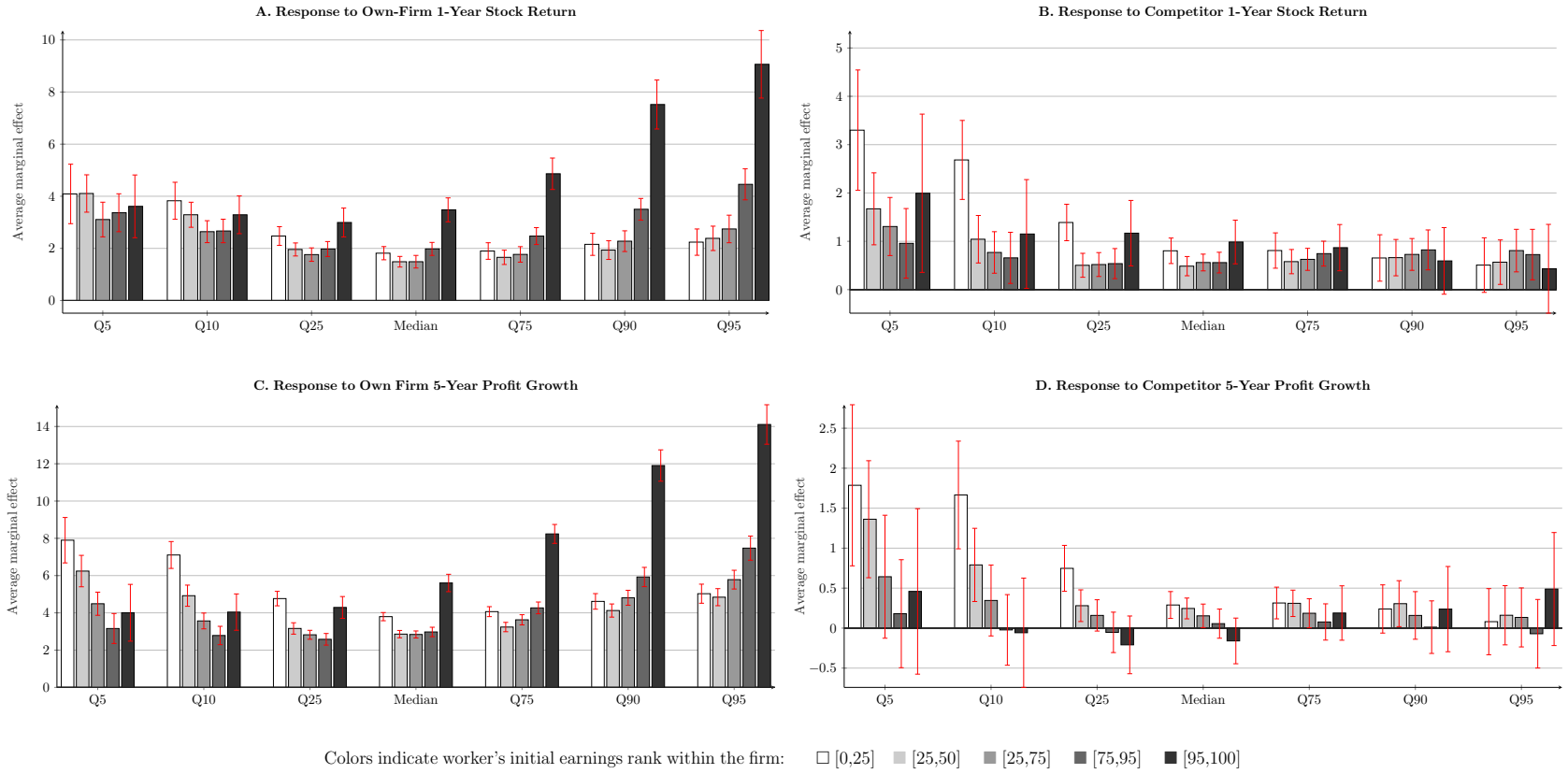


Note: This figure extends the analysis in Figure 8 to workers of all levels of income in the data. Specifically, we compare the innovation earnings exposure faced by continuing (Panel A) and moving (Panel B) workers. We plot the average marginal effects of firm—and competitor—innovation that are implied 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 one 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). In the data, movers/stayers are defined depending on whether they remain with the same firm within the next three years; see discussion in the main text for more details. 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 10: Earnings Risk and Own Firm Innovation: Heterogenous Innovations

Note: This figure examines whether the response of earnings growth percentiles to own firm innovation varies qualitatively with the type (novelty and/or process) of innovation. In Panel A, own firm innovation is separated into novel and less novel: a patent is classified as novel to the firm if its maximum similarity (cosine similarity, as in Kelly et al. (2020)) to the firm's own prior patents is sufficiently low (less than 0.5). In Panel B we examine process vs non-process innovation separately using the measure of Bena and Simintzi (2019). In Panel C we further decompose process patents into novel and less novel patents. For ease of comparison, marginal effects are scaled by the cross-sectional standard deviation of own firm innovation. The units on the vertical axis correspond to log points (times 100).

Figure 11: Worker Earnings Risk and Stock Returns or Profit Growth



Note: This figure illustrates that the response of worker earnings to innovation is qualitatively different from the response to other firm outcomes, such as profits or stock returns. 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 one 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).

Online Appendix

A Data Appendix

Here, we describe our data construction in more detail.

A.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.

A.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

[22] We access FOIA information for filings from 1999-present from the US Department of Labor's website. Information

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. The link between the form 5500 data and CCM data is based on Rauh and Stefanescu (2009) and Rauh, Stefanescu, and Zeldes (2020), 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.

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. The Supplementary Online Appendix contains more information on matched vs non-matched firms (EINs).

Note that, in section 2.1, we compute a number of nonparametric descriptive statistics for groups of workers sorted based on age, industry, and prior earnings rank. For that analysis, we include both public and privately held firms in the analysis, leveraging SIC code information from the MEF. Note that, for disclosure reasons, we can only obtain these estimates for cells with 100 or more workers. Thus, we omit any small cells with fewer than 100 matched workers from this analysis.

A.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:

$$Y_{i,t:t+h} \equiv w^i_{t,t+h} - w^i_{t-2,t}. \tag{A.1}$$

from 1990-1998 is taken from the Center for Retirement Research at Boston College University, available here.

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.1) 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 A.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:

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

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.

A.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:

$$\xi_j = (1-\bar{\pi})^{-1}\,\frac{1}{N_j}E[v_j|r_j]\;M_j. \tag{A.3}$$

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.3) 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, Hedge, and Marco, 2015). 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.3) 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,

$$\xi_{f,t}^{sm} = \sum_{j \in P_{f,t}} \xi_j. \tag{A.4}$$

In addition to the measures of innovation based on stock market reactions (A.4), 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

$$\xi_{f,t}^{cw} = \sum_{j \in P_{f,t}} \frac{1 + C_j}{1 + \bar{C}_j}, \tag{A.5}$$

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 as well as differences in citation patterns across technology classes. Both (A.4) and (A.5) 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,

$$A_{f,t}^{k} = \frac{\xi_{f,t}^{k}}{B_{ft}}, \qquad k \in \{sm, cw\}. \tag{A.6}$$

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,

$$A_{I\setminus f,t}^{k} = \frac{\sum_{f' \in I\setminus f} \xi_{f',t}^{k}}{\sum_{f' \in I\setminus f} B_{f't}}, \qquad k \in \{sm, cw\}. \tag{A.7}$$

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 A.4 contains more details.

To construct $A_{f,t}^{proc}$ and $A_{f,t}^{other}$, we use a similar procedure as equations (A.4) and (A.6). We create an estimate of the dollar amount of process innovation by the firm in year t as

$$\xi_{f,t}^{proc} = \sum_{j \in P_{f,t}} \theta_j \, \xi_j, \tag{A.8}$$

as well as the residual innovation $\xi_{f,t}^{other} = \xi_{f,t} - \xi_{f,t}^{proc}$. Similar to equation (A.6), 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, $\xi_{f,t}^{proc}/\xi_{f,t}$, is approximately 27.5%.

B Model Appendix

In this section, we provide additional technical details about the theoretical model discussed in the main text.

B.1 Firm and Wage Dynamics

In this section, we discuss how we construct the law of motion for product line-level profit and worker earnings dynamics from our theoretical model described in section 3 of the main text. Note that since each manager is paired with a single product line, the law of motion for product line and worker states are essentially identical whenever the worker remains employed at the same firm. We will place structure on the model so that all of the relevant state variables have a discrete support. However, worker and firm outcomes diverge once the worker separates. Accordingly, in this section, we will outline how to construct two separate generator matrices for the continuous time Markov chains which characterize worker and product line-level state dynamics respectively.

The firm which currently produces the good is called the "incumbent" and the firm which is not producing the good is called the "competitor." In the model, worker and firm dynamics can be summarized by the following four state variables, where we suppress subscripts for product lines:

1. κ_t: productivity ratio of producing (incumbent) firm and inactive, potential entrant (competitor) firm $\frac{q_t}{\tilde{q}_t}$
2. z_t: manager productivity of producing firm
3. $\lambda_{f,t}$: success rate of innovation for incumbent
4. $\lambda_{f',t}$: success rate of innovation for competitor

Because the model is set in continuous time, exogenous changes in state variables cannot happen simultaneously. The only exception is when the incumbent firm innovates, at which point they can draw a new value for $\log \kappa_t = \log \kappa_{t^-} + \log q_t - \log q_{t^-}$ and z_t at the same time.

To keep the dimension of the state space manageable, we restrict the support of κ and z to live on a discrete grid. In this context, the row of the generator matrix, which tracks the current state of $(\lambda_f, \lambda_{f'})$, z, and κ, will determine the probability of a given change in κ. Innovation intensity, given by

$\lambda_{f,t}$ and $\lambda_{f',t}$, follow independent two-point Markov processes, where each intensity has a probability of being high or low. Transitions between λ_L and λ_H states of $\lambda_{f,t}$ and $\lambda_{f',t}$ for each firm are given by the generator matrix

$$T = \begin{pmatrix} -\mu_H & \mu_H \\ \mu_L & -\mu_L \end{pmatrix}$$

These innovation intensities (λ_L, λ_H) and transition probabilities (μ_L, μ_H) are also free parameters in the model. When an incumbent firm innovates (where new innovations arrive at rate $\lambda_{f,t}$), its productivity increases, increasing $\log(\kappa)$ by the step size Δ_κ. Since means of κ and z are not separately identified, we normalize $\bar{z} = 1$. We impose an upper reflecting barrier on κ so that subsequent innovations do not improve the incumbent's competitive advantage (κ) when κ_{t^-} is already at its maximum value $\bar{\kappa}$.

For workers, an incumbent innovation also brings a change in z and a shock to her productivity which generates a nontrivial probability of being laid off. z evolves according to a discretized analog of a normal distribution such that $z_t \approx N(\rho_z z_{t^-} + \mu_z, \sigma_z)$. More precisely, we assume that $\log z$ has support on an equally spaced grid between $\log \bar{z} = 0$ and $\log z_{max}$. Analogous to Tauchen (1986), we set

$$P[z_t = z_j | z_{t^-}] = \Phi\left(\frac{\log z_j + \Delta_z/2 - \mu_z - \rho_z \log z_{t^-}}{\sigma_z}\right) - \Phi\left(\frac{\log z_j - \Delta_z/2 - \mu_z - \rho_z \log z_{t^-}}{\sigma_z}\right), \tag{A.9}$$

where Δ_z is the distance between adjacent points on the $\log z$ grid, and $\Phi(\cdot)$ is the cdf of the standard normal distribution. We set $P[z_t = \bar{z} | z_{t^-}] = \Phi\left(\frac{\Delta_z/2 - \mu_z - \rho_z \log z_{t^-}}{\sigma_z}\right)$ and $P[z_t = z_{max} | z_{t^-}] = \Phi\left(\frac{\log z_{max} - \Delta_z/2 - \mu_z - \rho_z \log z_{t^-}}{\sigma_z}\right)$, and we assume that the worker is terminated and replaced with a new one whenever z_t transitions from $z_{t^-} \geq \bar{z}$ downward to $\bar{z}$ following own firm innovation.

When a worker is replaced, she transitions into the pool of unemployed workers, from which she is rehired with intensity θ, at which point newly hired workers randomly match with firms who are hiring new workers each period. In other words, the probability that a worker, conditional on leaving unemployment, is hired by a firm with a particular $(\kappa, \lambda_f, \lambda_{f'})$ is proportional to the ergodic probability that a firm with these characteristics is hiring workers. (Recall that all newly hired workers begin with productivity $\bar{z}$.) New workers are hired by two types of firms: 1) firms replacing workers whose idiosyncratic z draws led them to be replaced, as described above, and 2) workers who are hired by new entrants who displace previous incumbents via the process of creative destruction which we outline next.

When a competitor innovates, transitions in $\log \kappa_t = \log \kappa_{t^-} - (\log \tilde{q}_t - \log \tilde{q}_{t^-})$ can either be large or small. When there is a high innovation intensity $(\lambda_{f',t} = \lambda_H)$, conditional on a successful innovation, the potential entrant has a chance to make large process improvements with probability p, which increase $log(\kappa)$ by a large amount, and small process improvements with probability $1 - p$, which decrease the competitive advantage of the current incumbent $log(\kappa)$ by Δ_κ. For simplicity, we assume that small competitor innovations cannot change the identity of the incumbent firm and that $\log \kappa$ cannot decline below a fixed threshold $\log \underline{\kappa} = 2\Delta_\kappa$ and also ensures that profits and wages are both positive.

In contrast to small innovations, large innovations always displace the incumbent, at which point the identity of the leading producer switches, and the current manager is fired. We assume that the

competitor draws a new κ from a distribution which is truncated below by z_{t^-} to approximate

$$\log(\kappa_t) \approx N(\log(\bar{\kappa}) - \log(\kappa_{t^-}), \sigma_\kappa^2) \quad | \quad log(\kappa_t) > z_{t^-},$$

where $\bar{\kappa}$ is the maximum value of κ on the grid. The truncation ensures that the competitor will always be able to earn higher profits than the current incumbent, and the conditional mean expression ensures that markups of the potential entrant are lower when its prior productivity ($\tilde{q}_{t^-}$) was substantially lower than the incumbent's (q_{t^-}) before the large innovation takes place. As before, we use a Tauchen-type formula to characterize the conditional probabilities associated with each value of κ_t, except that we truncate any values of κ_t which are lower than z_{t^-}. At the same time that an incumbent is displaced, the relevant Poisson intensities in the state space change to reflect the fact that the identity of the leading producer has changed. Thus, following a successful large innovation, $\lambda_{f,t} = \lambda_{f',t^-}$ and $\lambda_{f',t} = \lambda_{f,t^-}$.

We calibrate the model using 21 empirical quantiles, estimated on the subset of the data with workers whose initial wages are above the 95th percentile:

1. 7 quantiles for stationary wage growth rate for workers with initial wages
2. 7 quantiles for the marginal effect on wage growth rate of own firm innovation rate
3. 7 quantiles for the marginal effect on wage growth rate of competitor firm innovation rate

To solve the model, we fix the number of grid points over each variable, and the unemployment rate, and iterate on the free parameters until convergence. There are 10 free parameters in the model:

1. $\Delta\kappa^l$, the step size of $log(\kappa)$ conditional on a large innovation as an integer multiple of the small innovation step
2. μ_z, the expectation of $log(z)$ increment draw
3. σ_z, the standard deviation of $log(z)$ increment draws
4. σ_κ, the standard deviation of $log(\kappa)$ draws following large innovations of new entrants
5. λ_H, the high innovation intensity
6. λ_L, the low innovation intensity
7. μ_L, the intensity of transitioning from high to low innovation rate
8. μ_H, the intensity of transitioning from low to high innovation rate
9. p, the probability of large innovation when competitor successfully innovates
10. ρ_z, persistence coefficient for $log(z)$ draws conditional on own firm innovation

We fix some parameters, rather than estimating them.

1. Δ_κ, the step size of $\log(\kappa)$ conditional on a small innovation, is set to 0.0006
2. The upper bound for the $\log(z)$ grid is $0.99\log\bar{\kappa}$

We use 12 gridpoints for z, 87 gridpoints for κ, and have 2 levels for λ (both for the incumbent firm and the competitor firm). At each iteration, we take the previous values, and calculate the generator matrix through the rules described above. The generator matrix has rows corresponding to the current location of the state vector and columns corresponding to all possible transitions along the grid. We store transition intensities according to the conditional distributions and laws of motion discussed above, where we have one generator matrix for product-line level dynamics and another for worker-level wage dynamics. Once we have the generator matrix, we calculate the stationary distribution of the state vector as the

eigenvector associated with the generator matrix. After normalizing the stationary distribution to sum to 1, we calculate the wages implied by the model at each point. Since the level of wages plays little role in our setting, we normalize $\beta = 1$, which gives the wage at each state $w = (1 - \exp(\bar{z}))/\exp(z'\kappa')$.

When computing income growth rates, in the case where the worker is unemployed, we provide unemployment benefits equal to 10% of the smallest wage otherwise available to the workers. We calibrate the model by matching quantiles of five-year growth rates, and to get equivalent quantities from the generator matrix we add an identity matrix back to the generator matrix multiplied by some discrete timestep. In the calibration, we set the timestep Δ_t equal to $\frac{1}{12}$. Call the generator matrix Λ. Then the transition matrix over the discrete interval is approximated by $I + \Delta_t \Lambda$. Since we are calibrating the model against 5 year wage growth numbers, we update the transition matrix by multiplying it by itself $\frac{5}{\Delta_t}$ times, in this case 60.

Once we have the 5 year transition matrix, we compute the quantiles of the growth rates implied by the model. We estimate growth rates for the 5%, 10%, 25%, 50%, 75%, 90%, and 95% quantiles of wages, using both the stationary distribution, the distribution where the competitor successfully innovates, and where the worker's own firm successfully innovates. Quantiles are estimated by calculating the CDF of the growth rates for each distribution, since we have to weight the growth rate by the density at each point of the wage distribution. These are used directly for the baseline stationary distribution, but for each shock we instead match the marginal effect of the shock on the quantile of wage growth, scaling each by the standard deviation of the shock. Since the distribution of λ is given by a binomial distribution for the probability the competitor innovates and the probability the incumbent innovates, the variance of a shock is $\frac{\sqrt{\mu_H \mu_L}}{\mu_H + \mu_L}$. For industry innovation, which corresponds to a change in the average innovation rate for the product line, the relevant scaling factor for a one standard deviation change is $\frac{\sqrt{\mu_H \mu_L}}{2(\mu_H + \mu_L)}$, since incumbent and potential entrant's innovation intensities are independent.

B.2 Consumption-Savings problem

In this section, we provide additional details about the manner in which we solve the model. Since there is no aggregate risk in the model, our assumption that only a risk-free asset is tradable is without loss of generality. For simplicity we consider a model in which agents are infinitely lived. However, assuming that agents have a constant probability of death and no utility over bequests, ρ can alternatively be interpreted as the sum of an agent's rate of time preference parameter and the instantaneous hazard rate over bequests.

As is noted in the main text, agents in our model solve a fairly standard incomplete markets problem with post-tax income that follows an exogenous Markov chain. Let $V(\alpha, i)$ denote the value function given current wealth and that the agent is in income state i. As is well known, its solution satisfies the system of HJB equations for all $i = 1, \ldots, K$

$$0 = \max_{c \leq y(i) \text{ if } \alpha = 0} u(c) - \rho V(\alpha, i) + V_\alpha(\alpha, i)[\alpha \cdot r + y(i) - c] + \sum_{j \neq i} \psi_{i,j}[V(\alpha, j) - V(\alpha, i)], \quad \text{(A.10)}$$

where $\psi_{i,j}$ is the Poisson intensity of transitioning from state i to state j. If we recall that the diagonal

element of the generator matrix is equal to -1 times the sum of the off-diagonal elements within the same row, we note that this can be written in matrix notation as

$$\rho V(\alpha, i) = \max_c u(c(\alpha, i)) + V_\alpha(\alpha, i)[\alpha \cdot r + y(i) - c] + \Psi_i V(\alpha), \tag{A.11}$$

subject to

$$c \leq y(i) \quad \text{if} \quad \alpha = 0 \tag{A.12}$$

where Ψ_i is the i^{th} row of the generator matrix and $V(\alpha) \equiv [V(\alpha, 1), \ldots, V(\alpha, K)]'$. Even more compactly, the following system of equations must hold for a given level of assets α

$$\rho V(\alpha) = \max_c u(c) + V_\alpha(\alpha)[\alpha \cdot r + y - c] + \Lambda V(\alpha), \qquad c \leq y \quad \text{if} \quad \alpha = 0 \tag{A.13}$$

where Λ is the full generator matrix.

As is standard, when the borrowing constraint does not bind, the optimal consumption level satisfies the FOC:

$$c(\alpha, i)^{-\gamma} = V_\alpha(\alpha, i) \iff c(\alpha, i) = V_\alpha(\alpha, i)^{-1/\gamma}. \tag{A.14}$$

When $\alpha = 0$, we need to check whether the implied optimal consumption level is feasible. When this is not the case, it is optimal to set $c = y$.

We solve the Bellman equations on a grid over assets: $\{\alpha_1 = 0, \alpha_2, \ldots, \alpha_M\}$, and use an upwinding scheme to approximate $V_\alpha(\alpha, i)$ on each point on the grid.

Next, suppose that we initialize the value function at a particular guess $V(\alpha, 0)$.[23] We can then solve the HJB of the time-dependent problem which is identical to (A.15) except that we add additional $\frac{\partial V}{\partial t}$ term to the LHS (the term is added since we are iterating *backwards* when we do value function iteration):

$$\rho V(\alpha, t) + V_t(\alpha, t) = \max_c u(c) + V_\alpha(\alpha, t)[\alpha \cdot r + y - c] + \Psi V(\alpha, t), \qquad c \leq y \text{ if } \alpha = 0 \tag{A.15}$$

and we will use a finite difference approach to approximate this time derivative

$$\rho V(\alpha, t) + \frac{V(\alpha, t + \Delta) + V(\alpha, t)}{\Delta} \approx \rho V(\alpha, t) + V_t(\alpha, t) \tag{A.16}$$

where the time step Δ is a free parameter (higher Δ implies faster convergence at the cost of stability). Given the high dimensionality of the problem, we use an explicit scheme to update the value function. In this case, we can characterize the system of equations as follows for asset level α_i:

$$\rho V(\alpha_i, n + 1) + \frac{V(\alpha_i, n + 1) - V(\alpha_i, n)}{\Delta} \approx u(c_{i,n}) + V_\alpha(\alpha_i, n)[\alpha \cdot r + y - c_{i,n}] + \Psi V(\alpha_i, n), \tag{A.17}$$

where $c_{i,n}$ denotes the solution for optimal consumption that obtains at each one of the income states for asset level i using the guess of the value function based on the previous guess of the value function

[23] Specifically, we initialize the value function with the utility associated with the constant consumption stream the agent could maintain forever if income stayed fixed at the current state, given the current level of assets.

$V(\alpha, n)$. We iterate on equation (A.17) until we reach convergence.

Once we solve the value function, it is straightforward to compute the ergodic distribution of assets for each income level by solving the Kolmogorov forward (Fokker-Plack) equation (see, e.g., Achdou et al., 2020, for more details). Given our upwinding approximation of partial derivatives, we immediately obtain an approximation of the infintessimal generator matrix of a Poisson approximation of the joint distribution of income states and assets. The stationary distribution obtains as the eigenvector associated with the zero eigenvalue of the transpose of this matrix, rescaled to be positive and sum to unity.

C Econometric Methodology

Here, we relegate details of our estimation methodology.

In our analysis, we use a method for estimating multiple conditional quantiles recently developed in Schmidt and Zhu (2016). This method, which we briefly describe next, 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.

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

$$q(\alpha; x) \equiv \inf\{y \in \mathbb{R} \colon P[Y_{i,t} \leq y \;\mid\; X_{i,t} = x] \geq \alpha\}. \tag{A.18}$$

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:

$$q_j(x) = \begin{cases} x'\beta_0 & \text{if } j = j^* \\ x'\beta_0 - \sum_{k=j}^{j^*-1} \exp(x'\beta_k) & \text{if } j < j^* \\ x'\beta_0 + \sum_{k=j^*+1}^{j} \exp(x'\beta_{k-1}) & \text{if } j > j^* \end{cases}. \tag{A.19}$$

The econometric model in (A.19) 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).[24]

[24] 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.19) by iteratively applying a sequence of standard linear-in-parameters quantile regressions, beginning with the median and working toward the tails.

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

$$\beta_j = \frac{\partial}{\partial x}[\log(q_{j+1}(x) - q_j(x))], \tag{A.20}$$

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.

An advantage of our estimation methodology is that it results in highly tractable forms for average marginal effects,

$$E\left[\frac{\partial q_j(X_{i,t})}{\partial X_{i,t}}\right] = \begin{cases} \beta_0 & \text{if } j = j^* \\ \beta_0 - \sum_{k=j}^{j^*-1} E[\exp(X'_{i,t}\beta_k)]\beta_k & \text{if } j < j^* \\ \beta_0 + \sum_{k=j^*+1}^{j} E[\exp(X'_{i,t}\beta_{k-1})]\beta_k & \text{if } j > j^* \end{cases}. \tag{A.21}$$

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.

References

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2020). Income and wealth distribution in macroeconomics: A continuous-time approach. *Review of Economic Studies forthcoming*.

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. Hedge, and A. Marco (2015). What is the probability of receiving a US patent. *Yale JL & Tech. 17*, 203.

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

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.

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 (2020). Cost saving and the freezing of corporate pension plans. *Journal of Public Economics 188*, 104211.

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

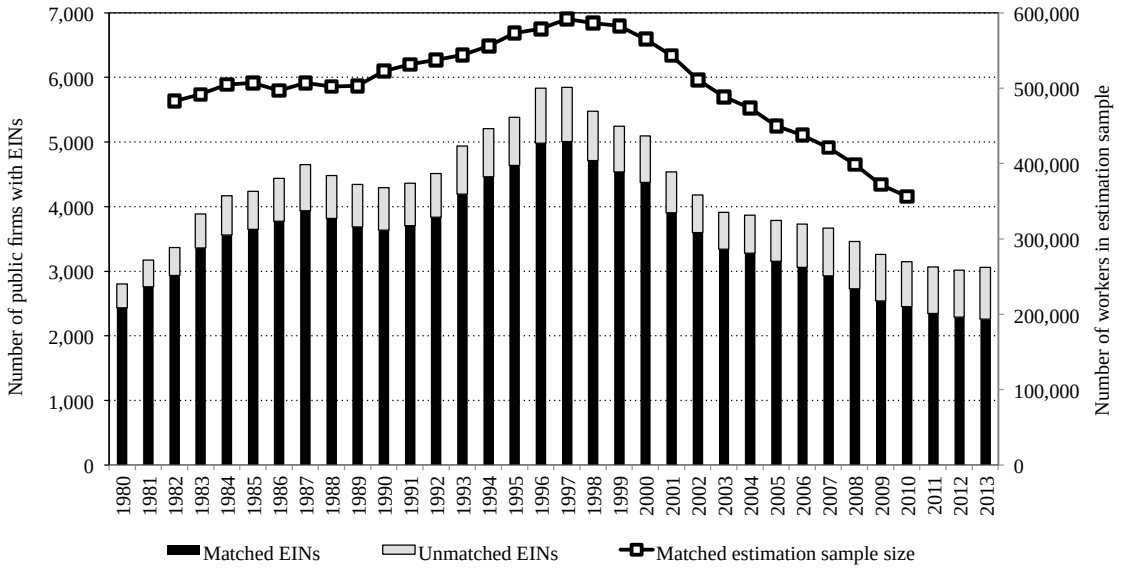
Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *The Quarterly Journal of Economics 134*(1), 1–50.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters 20*(2), 177–181.

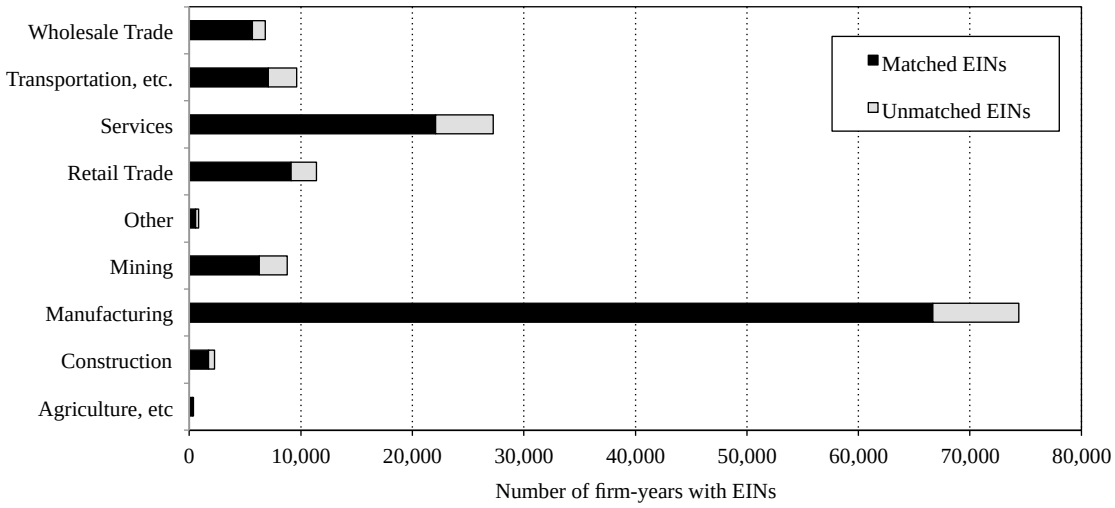
Appendix Tables and Figures

Figure A.1: Characteristics of the matched sample

Panel A: Number of observations by year

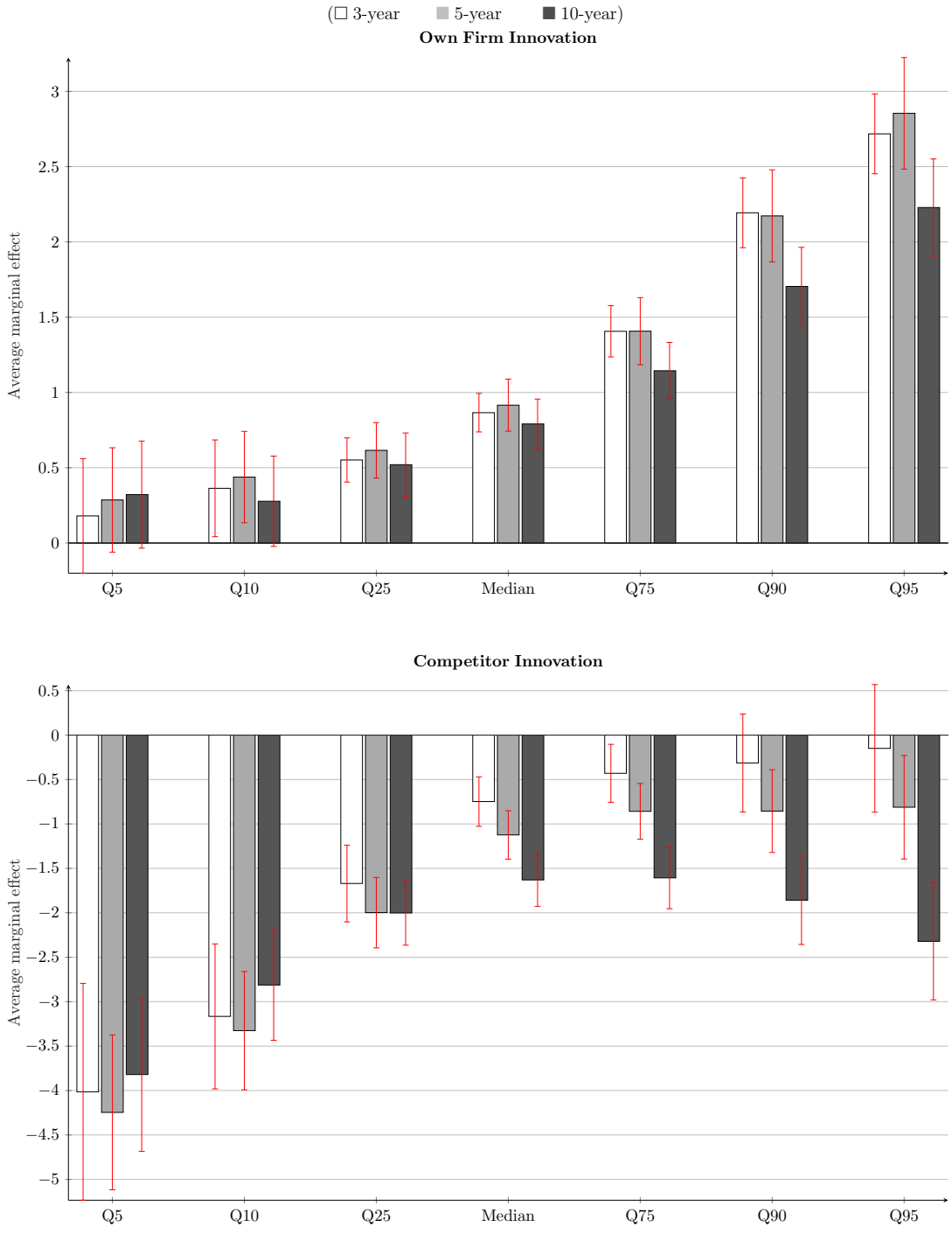


Panel B: Number of observations by major SIC sector



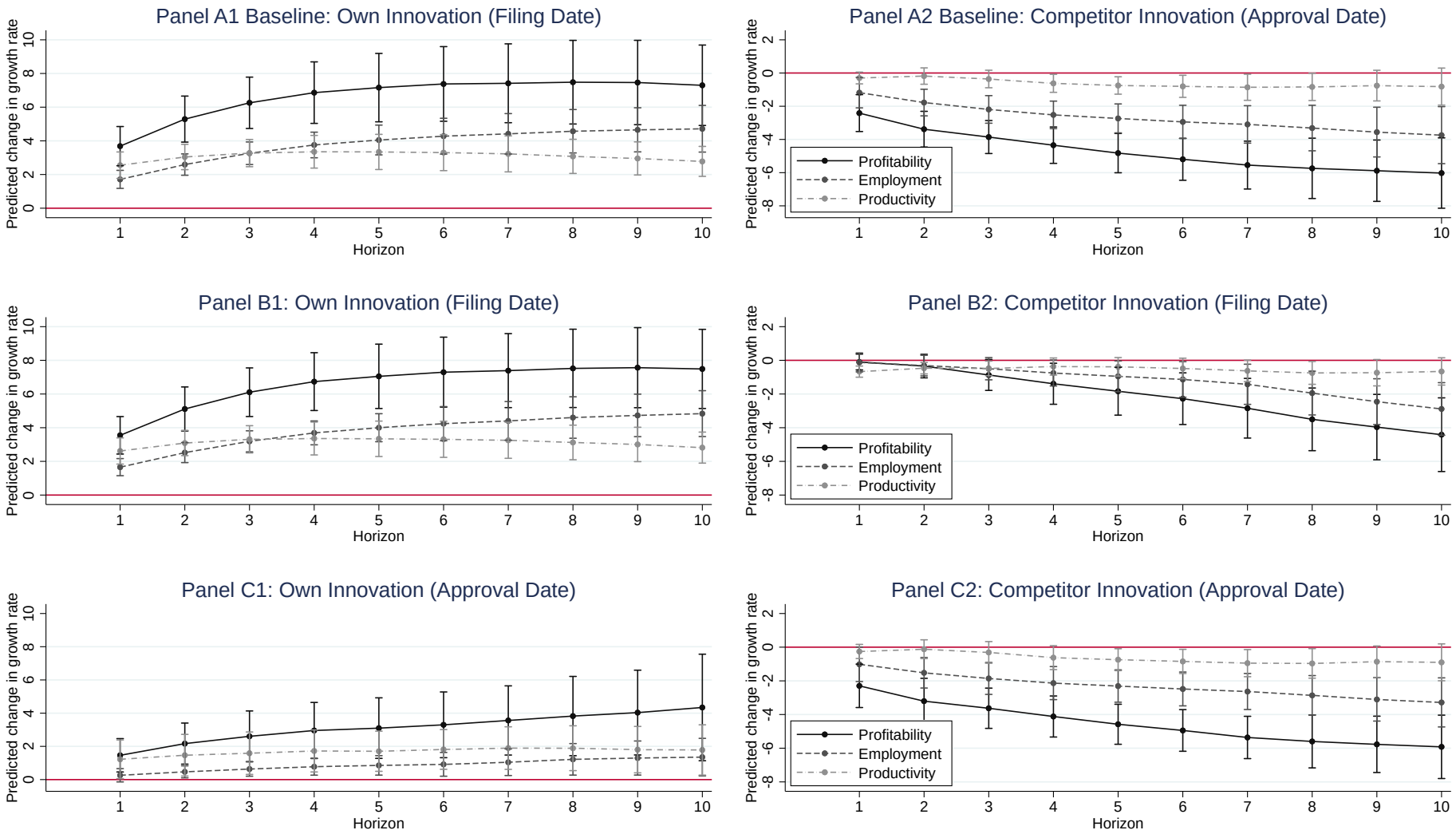
Note: The top plot provides counts of the number of public-firm years for which we can find matched W-2 earnings records in the SSA master earnings file, as well as the number of firm-years for which no earnings records could be found. We exclude firm-years for which no EIN is available. The bottom panel repeats the analysis by major SIC sector, where the SIC codes are taken from firms' financial statements.

Figure A.2: Earnings growth and innovation across horizons



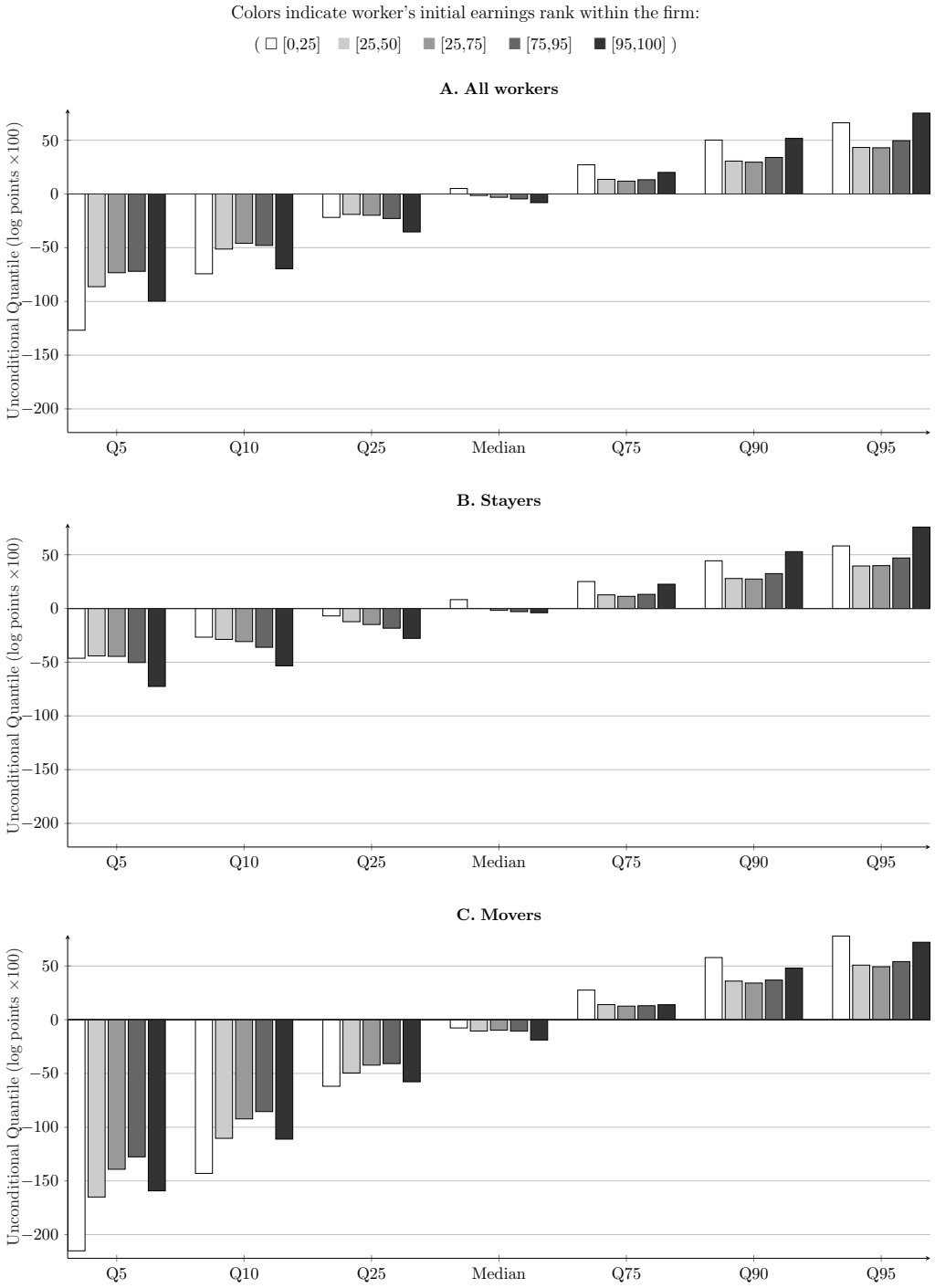
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) across different horizons (number of years of cumulative future earnings included in earnings growth), where estimates are scaled to correspond with a one standard deviation change in each variable of interest. The units on the vertical axis correspond to log points (times 100).

Figure A.3: Innovation and growth - Firm-level outcomes across horizons, varying timing conventions



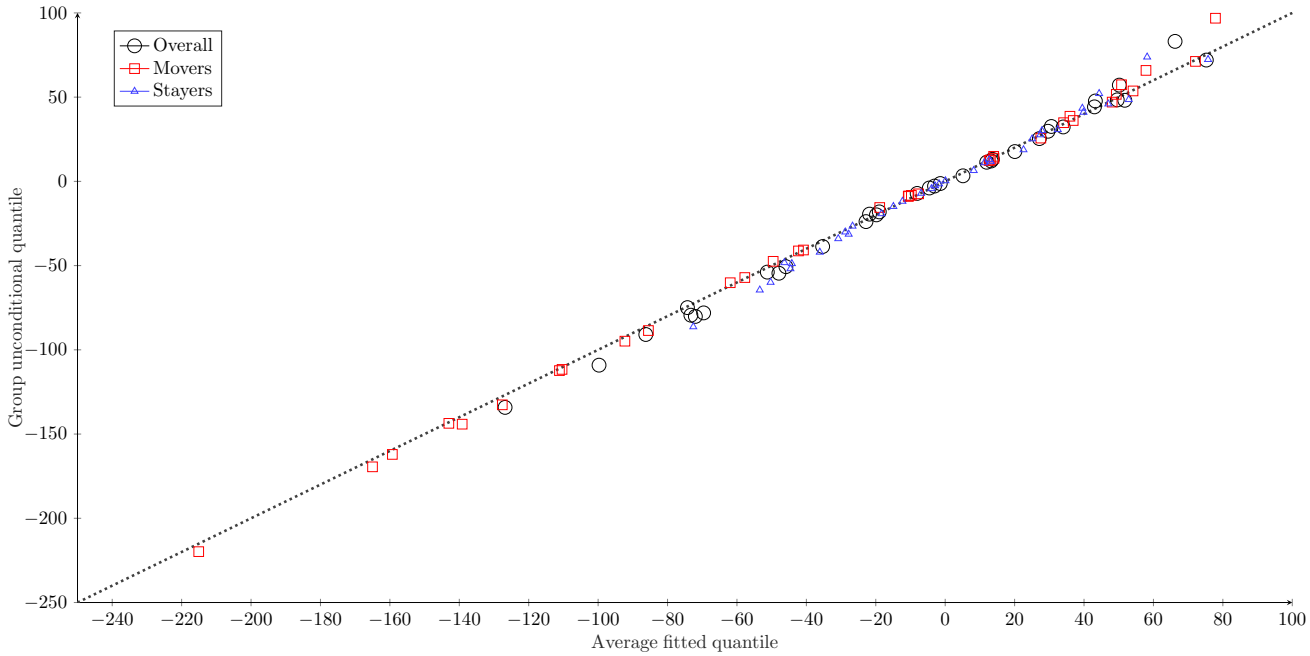
Note: Figure reports coefficient estimates of equation (11) for firm profits, employment and TFPR. The horizontal axis varies the horizon of the regression. Each dependent variable corresponds with a different line on the graph. Each specification relates firm growth to innovation by the firm (A_f, defined in equation (3) 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 (4)). Panels B and C run the same regressions, changing the timing convention of own and competitor innovation measures to use the filing and approval dates, respectively. 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.

Figure A.4: Percentiles of income growth: movers versus continuing workers



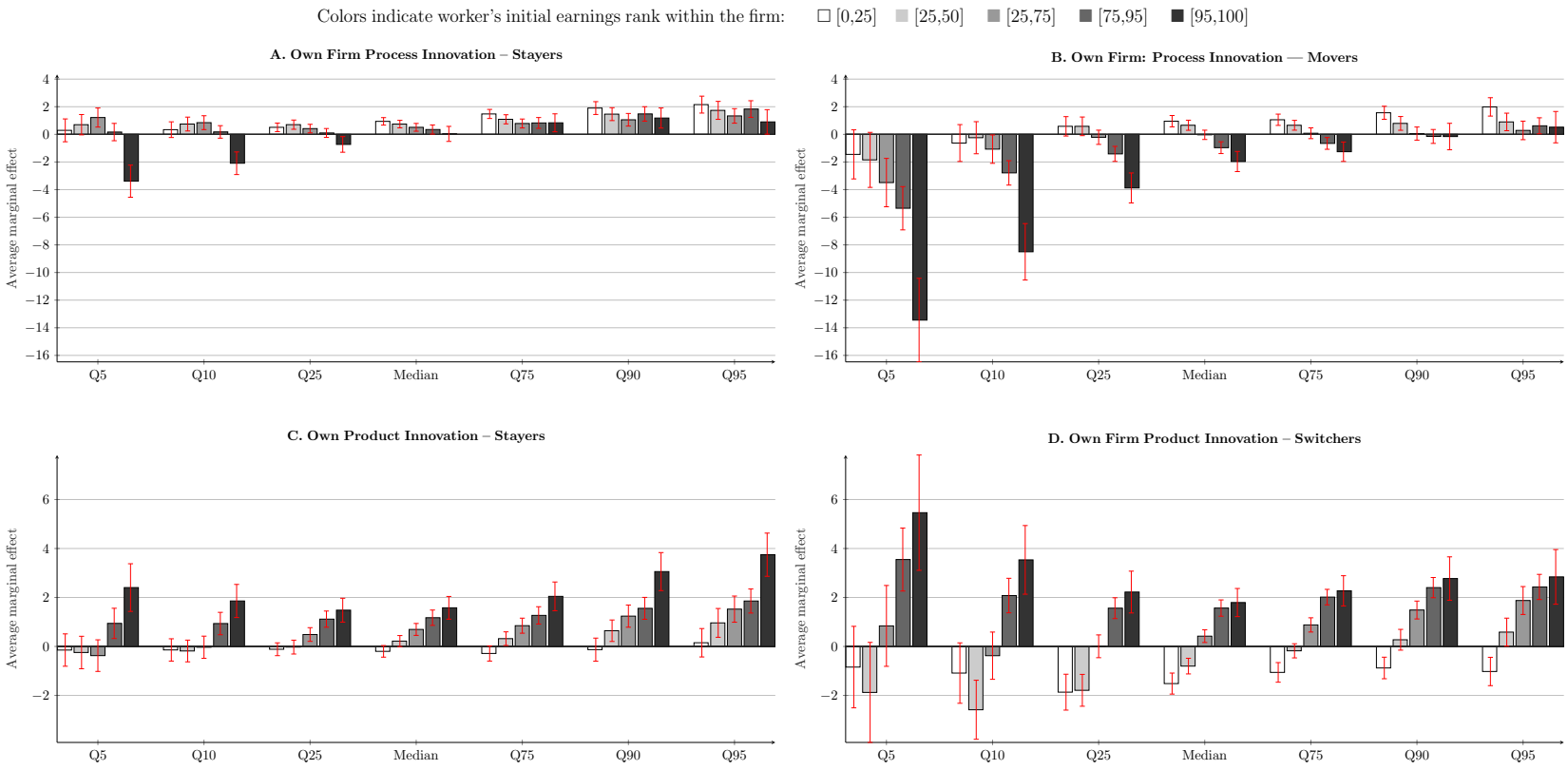
Note: Figure plots the distribution of 5-year earnings growth for workers of different earnings levels (earnings ranks) separately for workers that remain with the same firm after 5 years (stayers) and for those that do not (movers). The top panel plots the average fitted quantiles from the same specification as Figure 3. The middle panel presents coefficients from the same specification, estimated for the subsample of movers–workers who are employed at the firm in year t+3; while the bottom panel estimates the same model on workers who are not employed at the same firm in year t+3. These figures are based on average conditional quantiles from the estimated model. In general, the average conditional quantiles need not correspond to the unconditional quantiles for each group; however, in our case they do. Appendix Figure A.5 shows these average predicted quantiles are quite similar to the unconditional quantiles in the data.

Figure A.5: Unconditional quantiles versus average fitted quantiles by firm rank bin



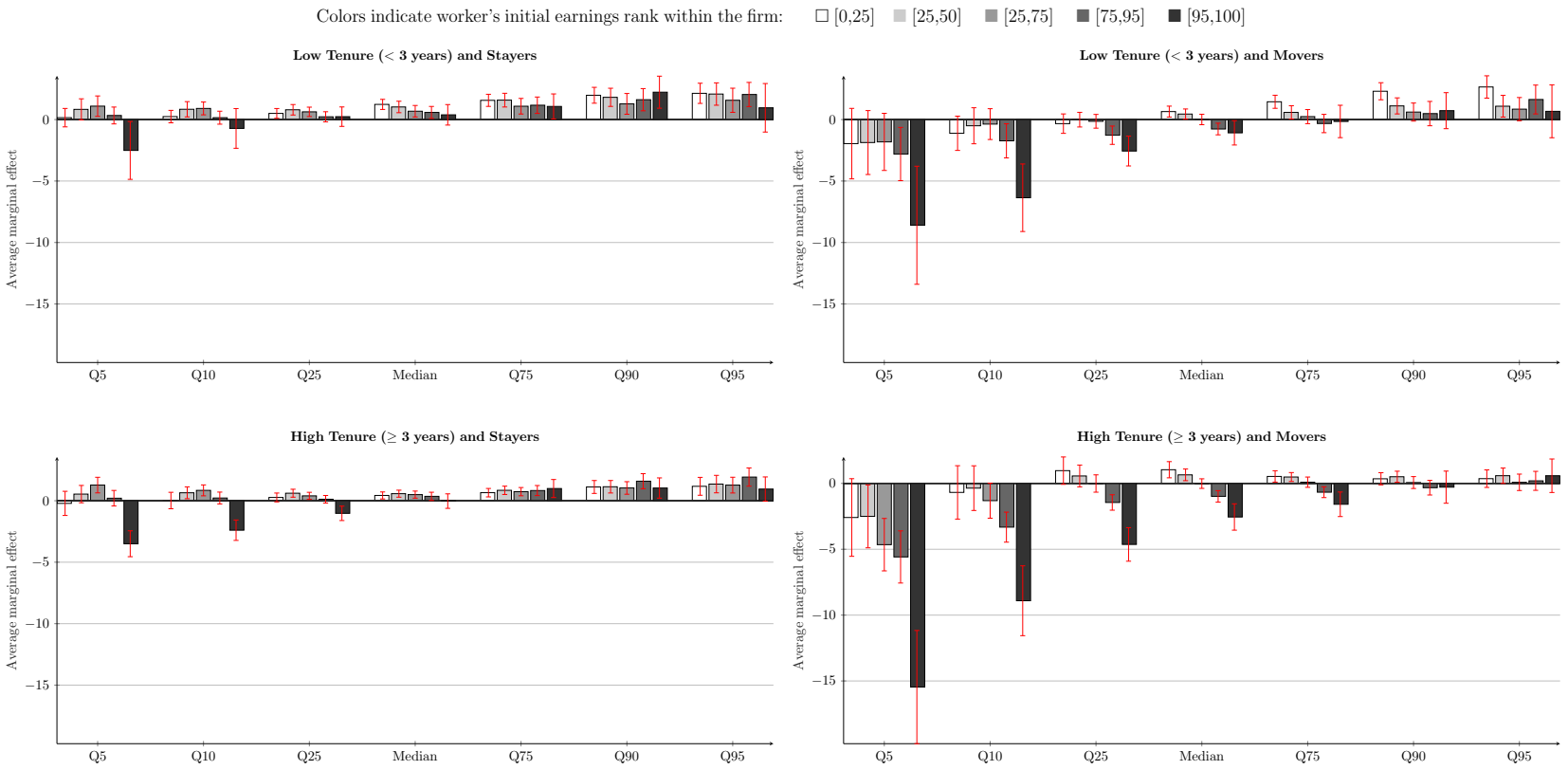
Note: Figure compares the average fitted quantiles plotted from Figure A.4 of the distribution of 5-year earnings growth for workers of different earnings levels (earnings ranks) with raw unconditional quantiles calculated for each group. In addition to the specification which is estimated for the full sample, we also repeat the exercise separately for the subsamples of workers that remain with the same firm after 5 years (stayers) and for those that do not (movers). Stayers are defined as workers who are employed at the firm in year t+3; while movers workers who are not employed at the same firm in year t+3.

Figure A.6: Earnings growth and own firm process/product innovation conditional on worker earnings levels: movers versus continuing workers



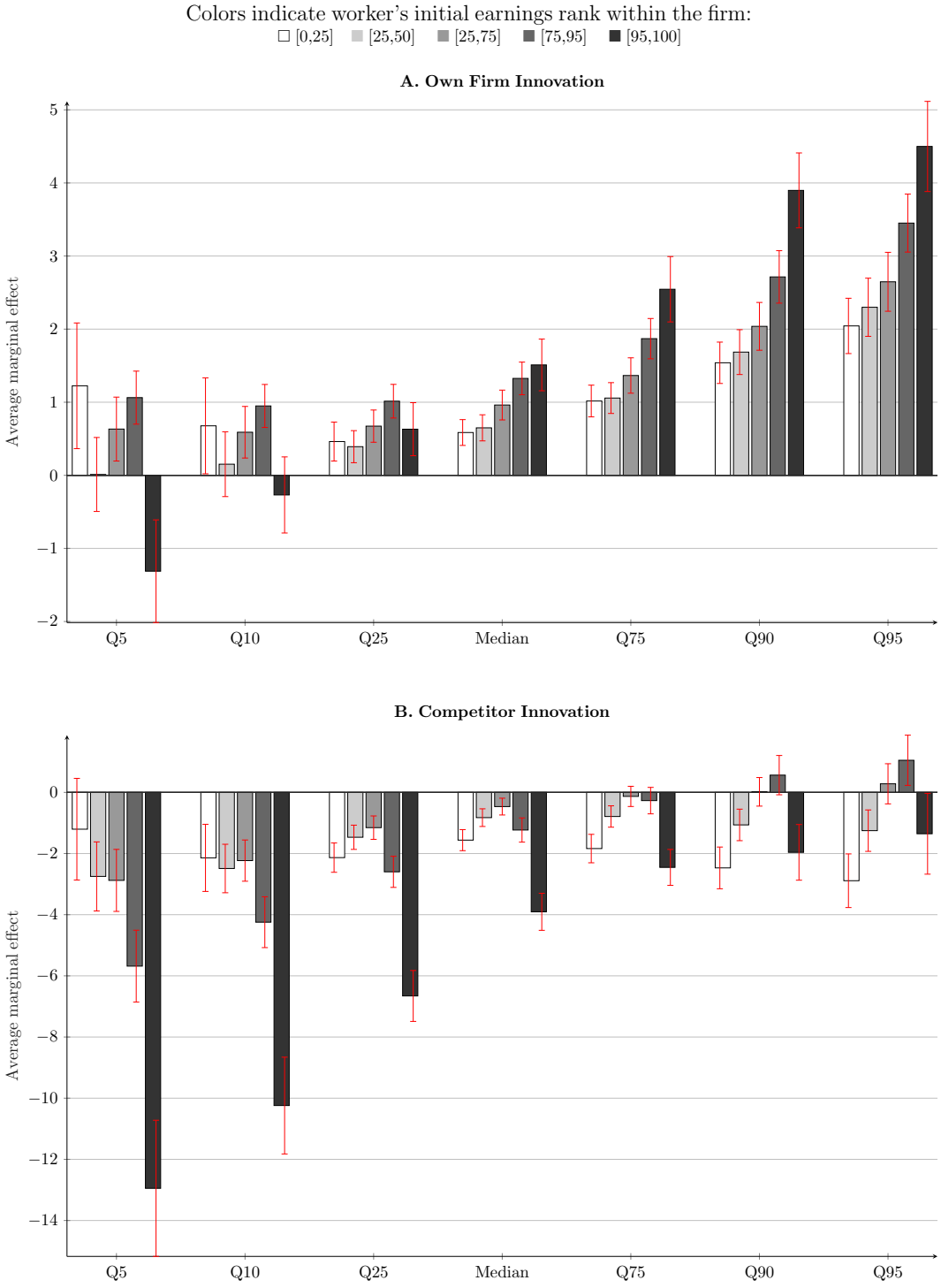
Note: Figure plots the average marginal effects of firm process/product innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. Estimates are standardized to correspond with one standard deviation effects for each variable 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-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.7: Earnings growth and own firm process innovation conditional on worker earnings levels and tenure: movers versus continuing workers



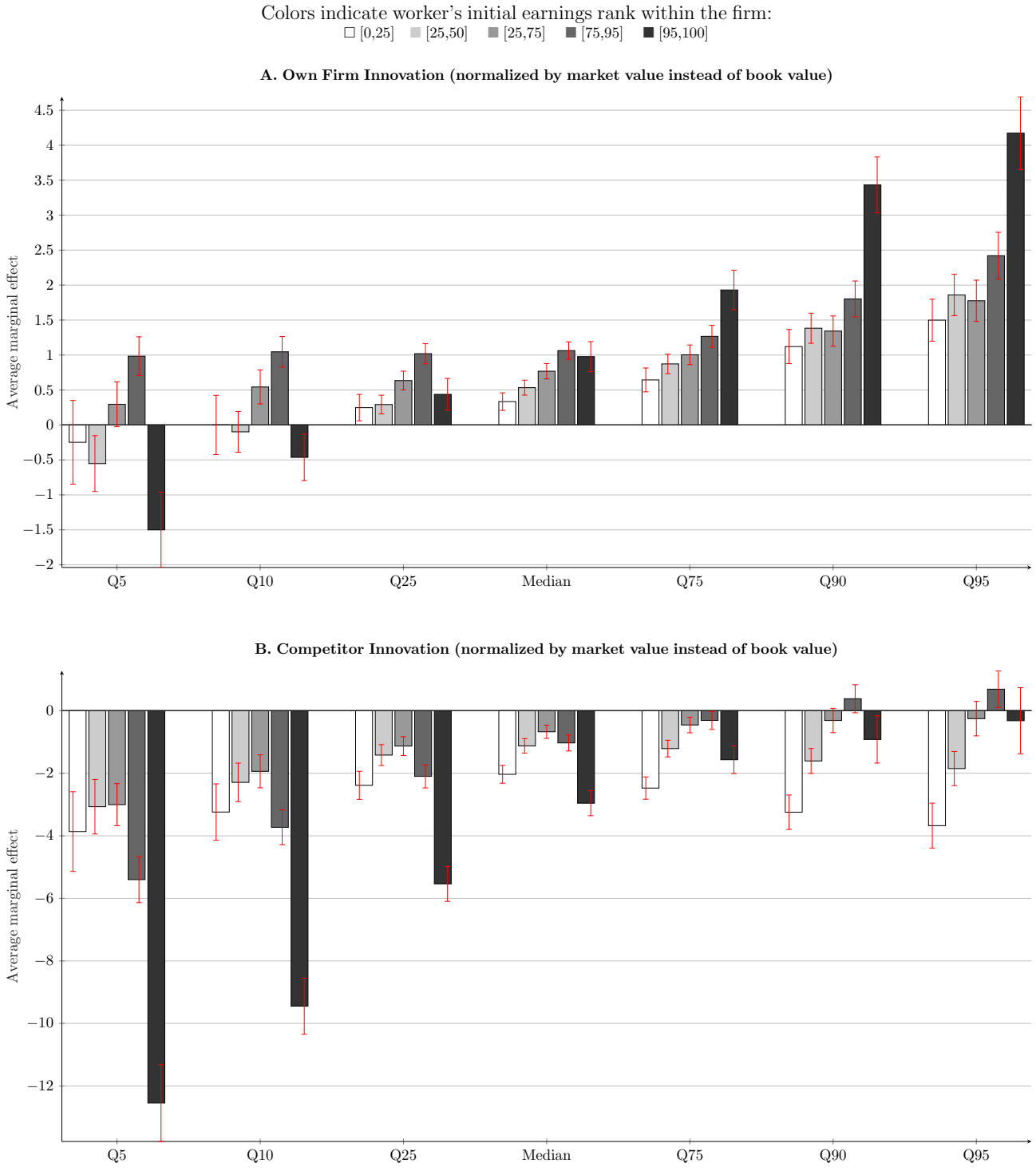
Note: Figure plots the average marginal effects of firm process innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels and for workers with different earnings levels and years of tenure with the firm. Workers are sorted into two groups based upon whether they have less than 3 years or greater than or equal to 3 years of tenure, and we allow for separate coefficients for each tenure × lagged earnings bin. Estimates are standardized to correspond with one standard deviation effects for 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-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.8: Earnings growth and innovation conditional on earnings levels - control for R&D spending



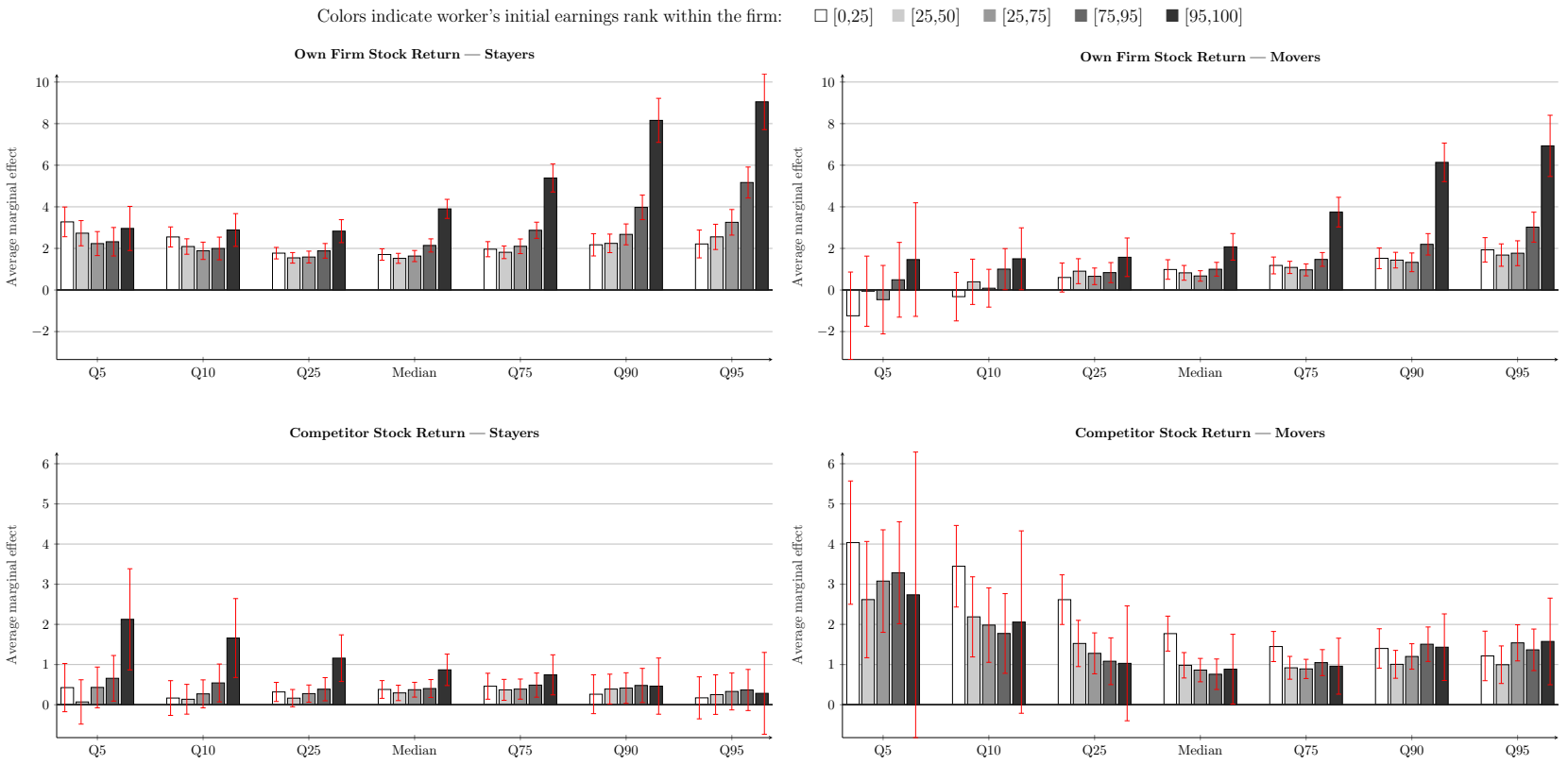
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels, where we additionally control for the ratio of R&D to assets in the regression (and drop firms with missing R&D data). Estimates are standardized to correspond with one standard deviation effects for each variable of interest. The worker earnings rank is defined net of deterministic life cycle effects. We focus on 5-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.9: Earnings growth and innovation conditional on earnings levels - alternative firm value scaling factor for innovation measure



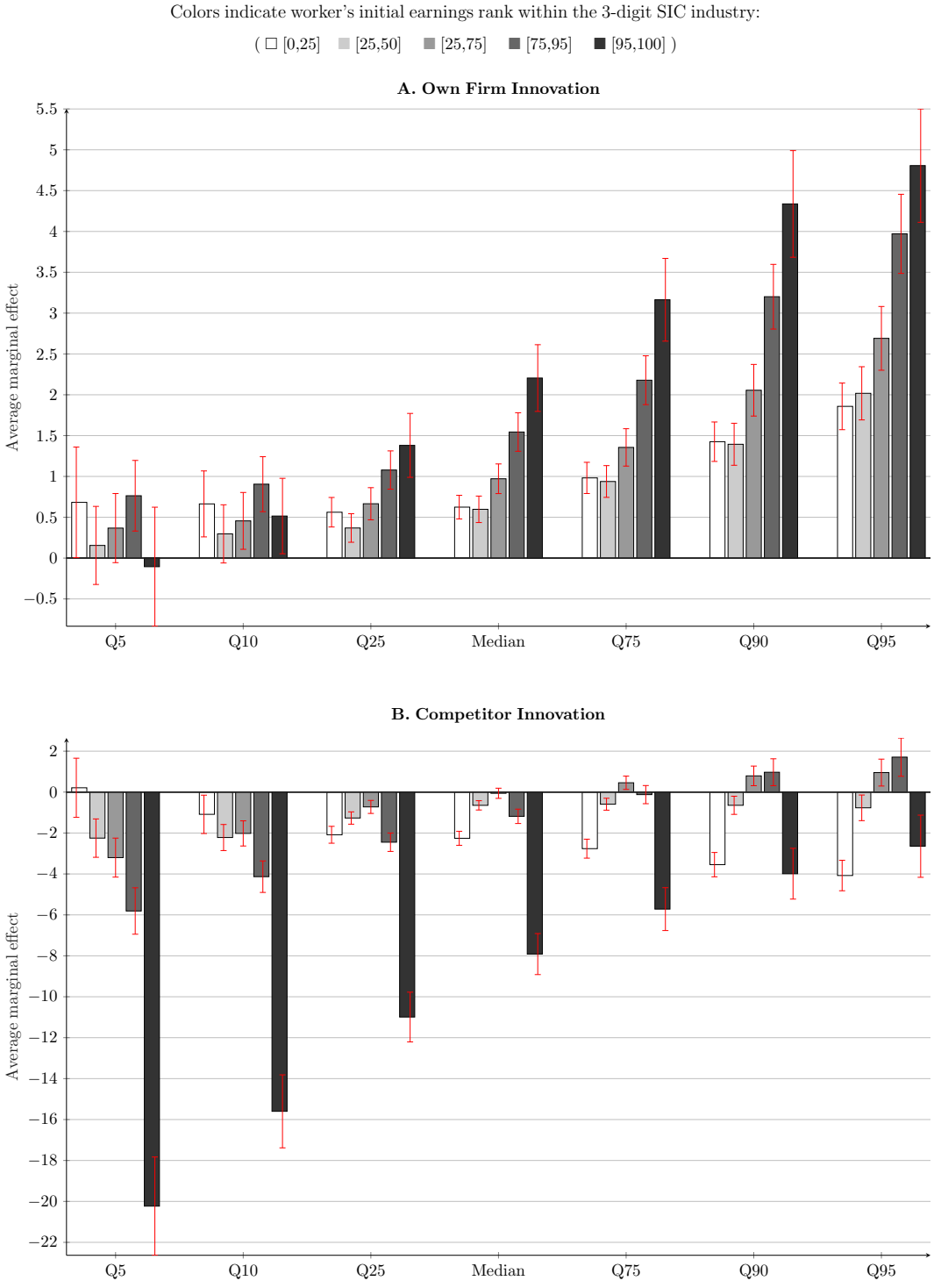
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. Estimates are standardized to correspond with one standard deviation effects for each variable of interest. The difference from the baseline specification is that our innovation measure is scaled by the market value, rather than the book value, of firm assets. The worker earnings rank is defined net of deterministic life cycle effects. We focus on 5-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.10: Earnings growth and stock returns: movers versus continuing workers



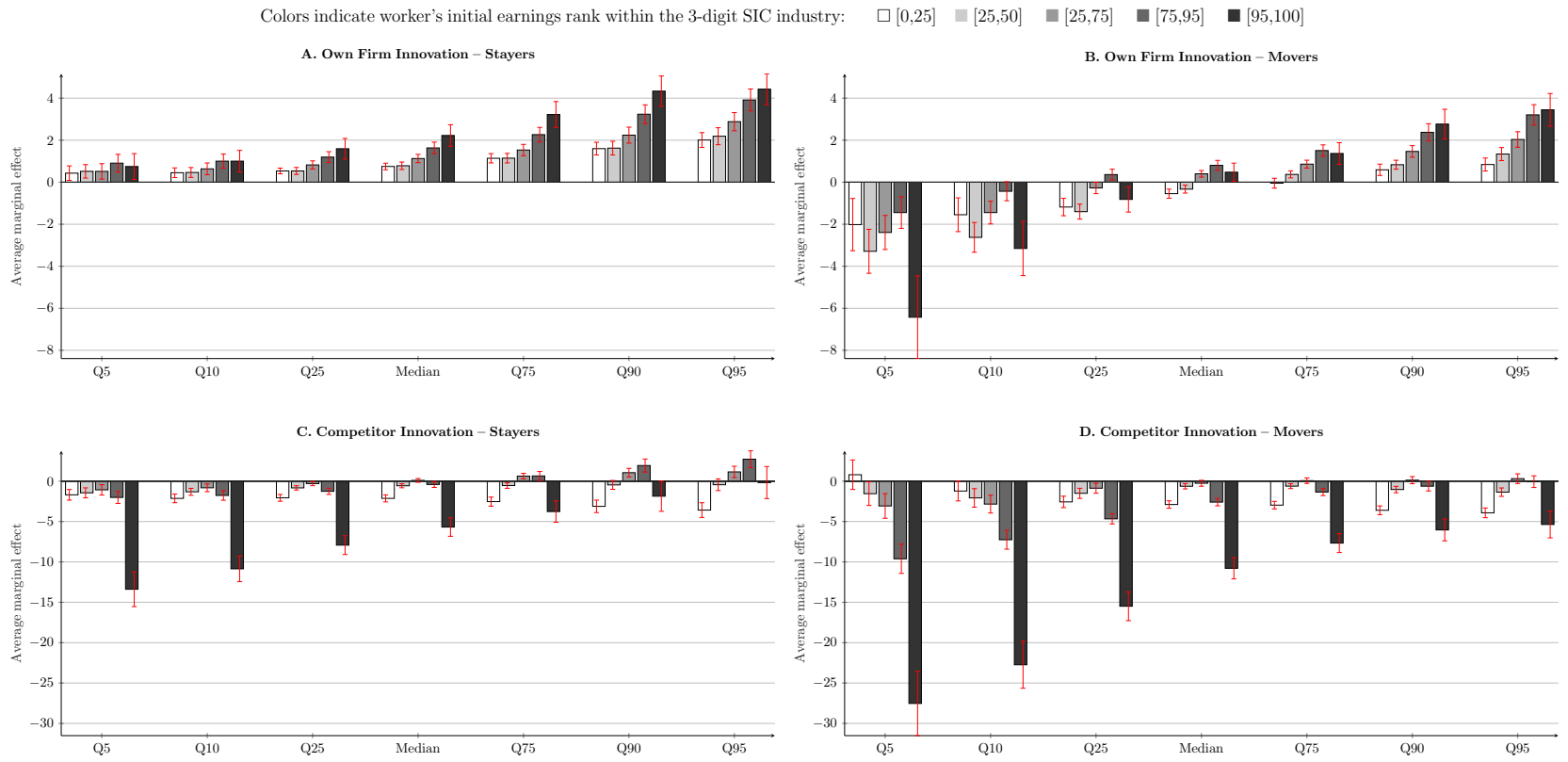
Note: Figure plots the average marginal effects of firm—and competitor—stock returns that are implied by the quantile regression estimates (analogous to equation (12) in the main text, except that we use own firm and competitor year t returns in place of the innovation measures) for workers with different earnings levels. Estimates are standardized to correspond with one standard deviation effects for 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 A.11: Earnings growth and innovation conditional on earnings levels – sort on industry income rank



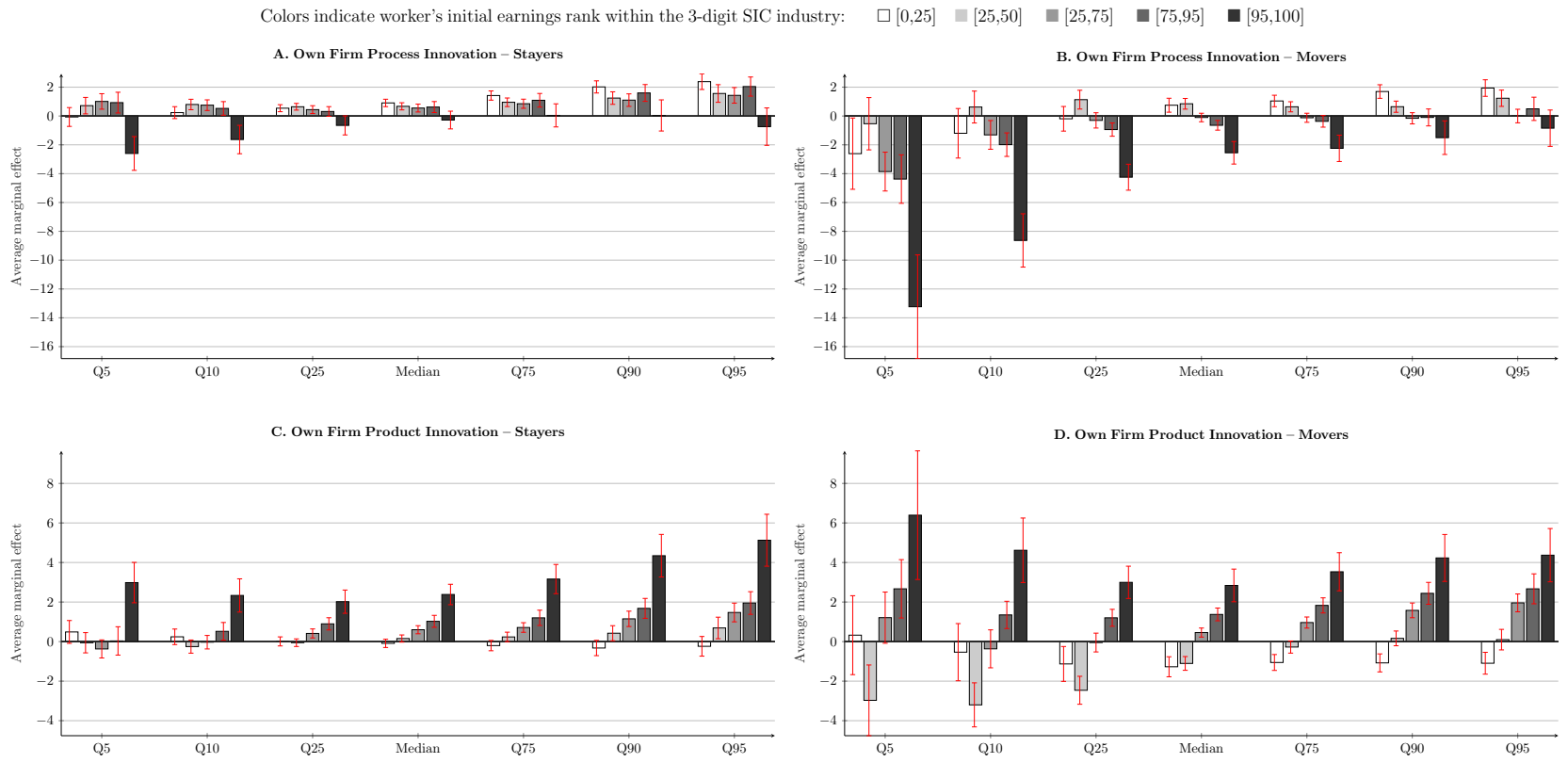
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels, where estimates are scaled to correspond with a one standard deviation change in each variable of interest. Whereas the baseline specification sorts on rank within the firm, here we compute ranks within the same 3-digit SIC industry. 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 A.12: Earnings growth and innovation: movers versus continuing workers – sort on industry income rank



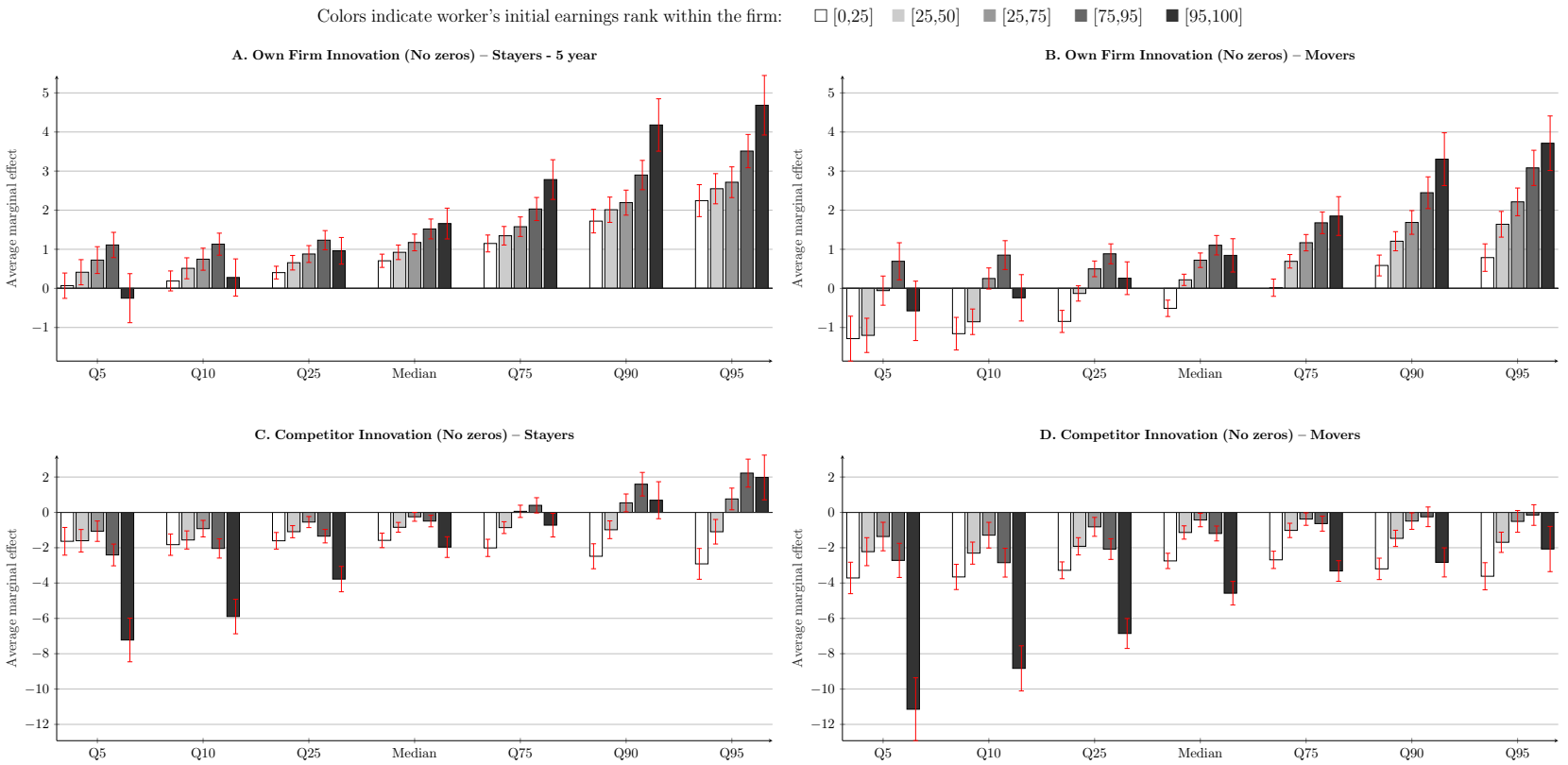
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels, where estimates are scaled to correspond with a one standard deviation change in each variable of interest. The equation is estimated separately for workers that remain with the firm (stayers) versus workers that leave the firm (switchers). Whereas the baseline specification sorts on rank within the firm, here we compute ranks within the same 3-digit SIC industry. 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 A.13: Earnings growth and process/product innovation: movers versus continuing workers – sort on industry income rank



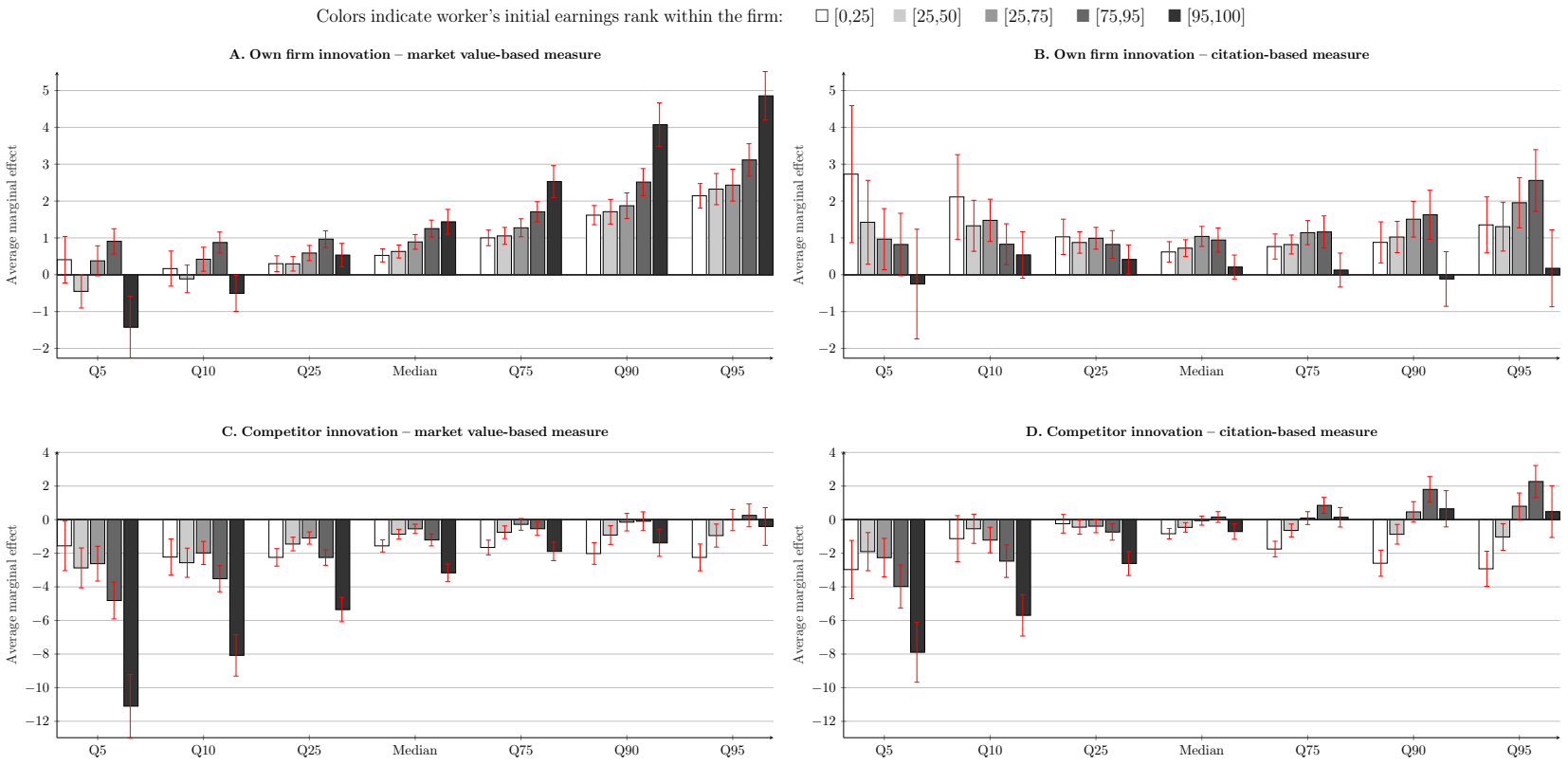
Note: Figure plots the average marginal effects of firm process/product innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels, where estimates are scaled to correspond with a one standard deviation change in each variable of interest. The equation is estimated separately for workers that remain with the firm (stayers) versus workers that leave the firm (switchers). Whereas the baseline specification sorts on rank within the firm, here we compute ranks within the same 3-digit SIC industry. 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 A.14: Earnings growth and innovation: movers versus continuing workers - exclude years with zero income obs



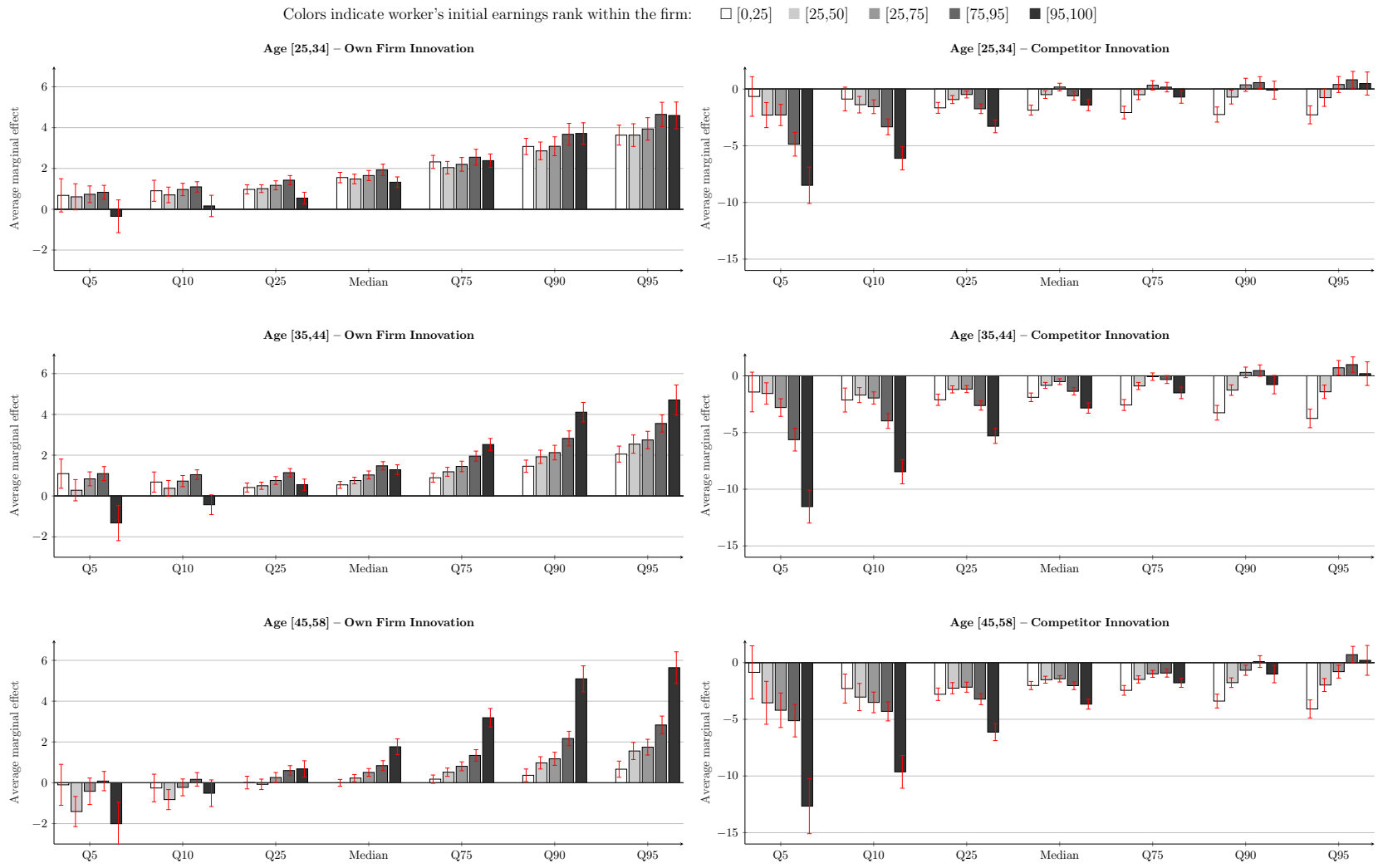
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels, where workers with any years with zero W-2 earnings are excluded from the estimation. Estimates are standardized to correspond with one standard deviation effects for 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-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.15: Earnings growth and innovation: valuable vs highly-cited patents



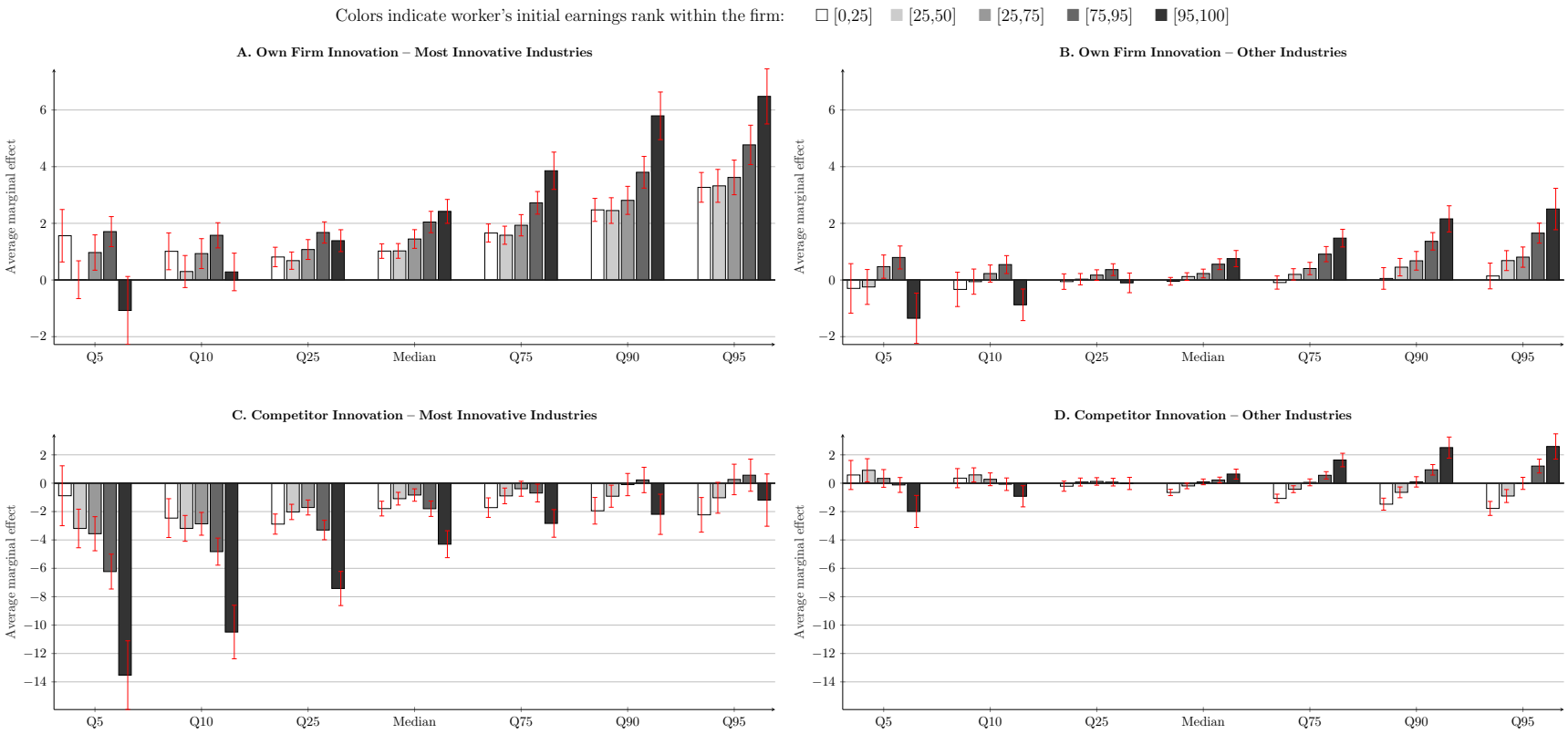
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. In addition to using our market value-based measures of own firm and competitor innovation ($A^{sm}_{f,t}$ and $A^{sm}_{I\setminus f,t}$, respectively), we additionally include their citation-based analogs: $A^{cw}_{f,t}$ and $A^{cw}_{I\setminus f,t}$. Estimates are standardized to correspond with one standard deviation effects for 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-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.16: Earnings growth and innovation conditional on age and earnings levels



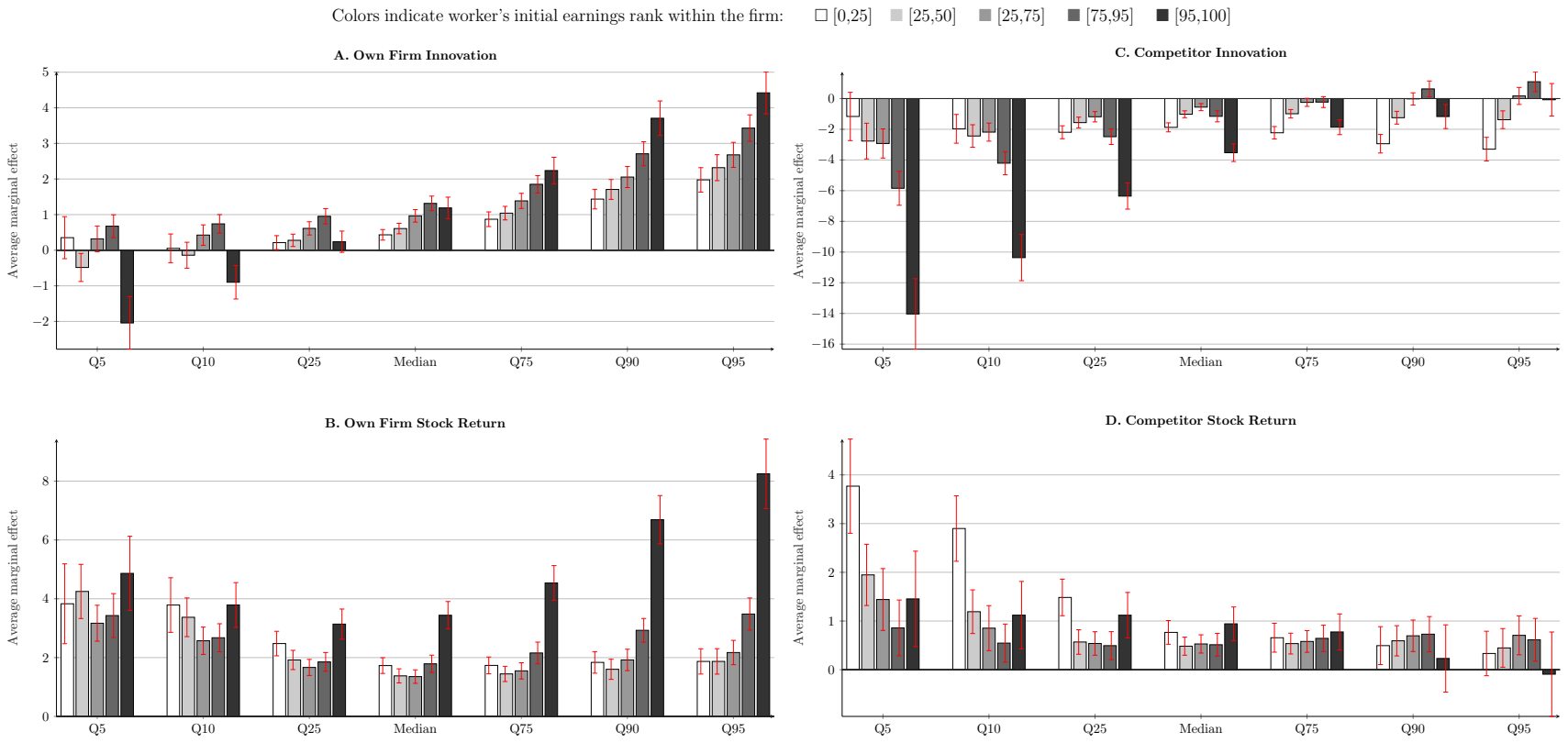
Note: Figure plots the average marginal effects of firm innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different ages and earnings levels. Estimates are standardized to correspond with one standard deviation effects for each variable 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 A.17: Earnings growth and innovation conditional on worker earnings levels: most innovative versus other industries



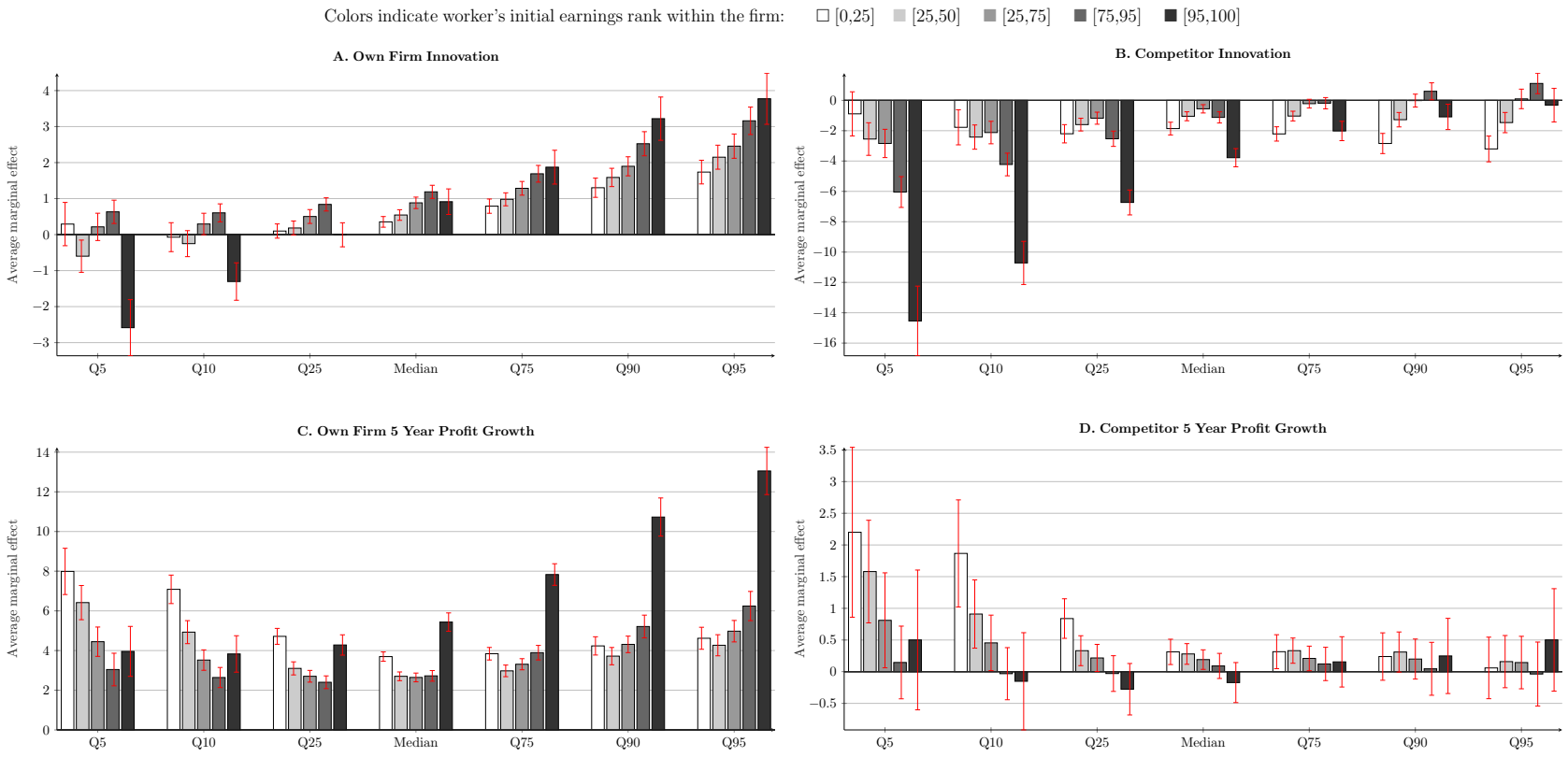
Note: Figure plots the average marginal effects of own firm and competitor innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. The equation is estimated separately for workers in industries which are in the top tercile of innovativeness, which is defined as the average across years of the ratio of the market value of patents in each year (aggregated across firms within an industry year) divided by lagged book assets. We choose the breakpoint so that 1/3 of 3-digit SIC codes (weighted by Compustat employment) are included in the high-tech category. Estimates are standardized to correspond with one standard deviation effects for each variable of interest. The worker earnings rank is defined net of deterministic life cycle effects. We focus on 5-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.18: Earnings growth and innovation: responses to own firm and competitor innovation, controlling for own/competitor stock returns



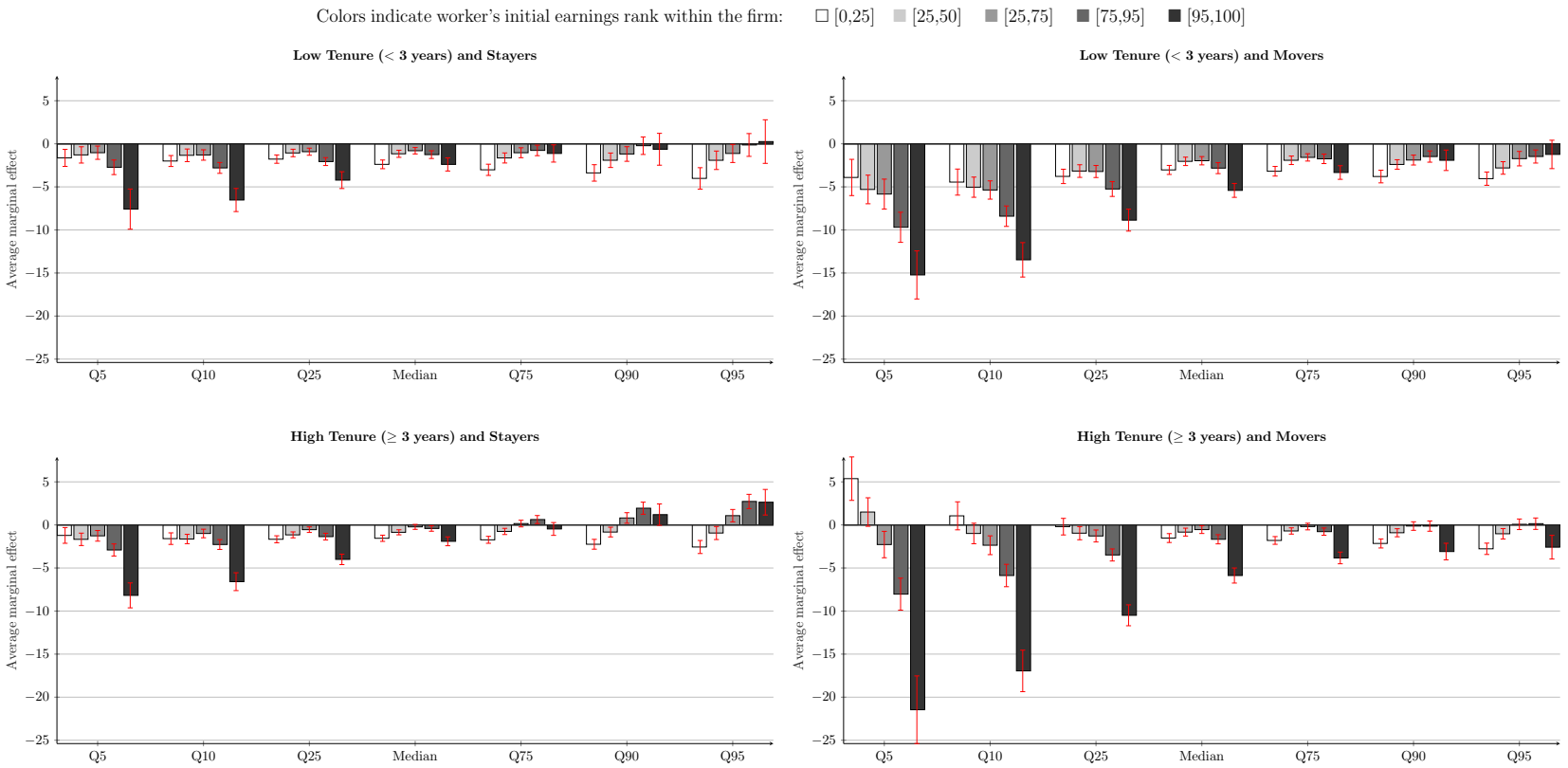
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. In addition to the own firm and competitor innovation measures from the baseline specification, we also include controls for own firm and competitor 5 year cumulative profit growth and also allow these coefficients to vary with a worker's rank within the firm. Estimates are standardized to correspond with one standard deviation effects for each variable of interest. The worker earnings rank is defined net of deterministic life cycle effects. We focus on 5-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.19: Earnings growth and innovation: responses to own firm and competitor innovation, controlling for own/competitor profit growth



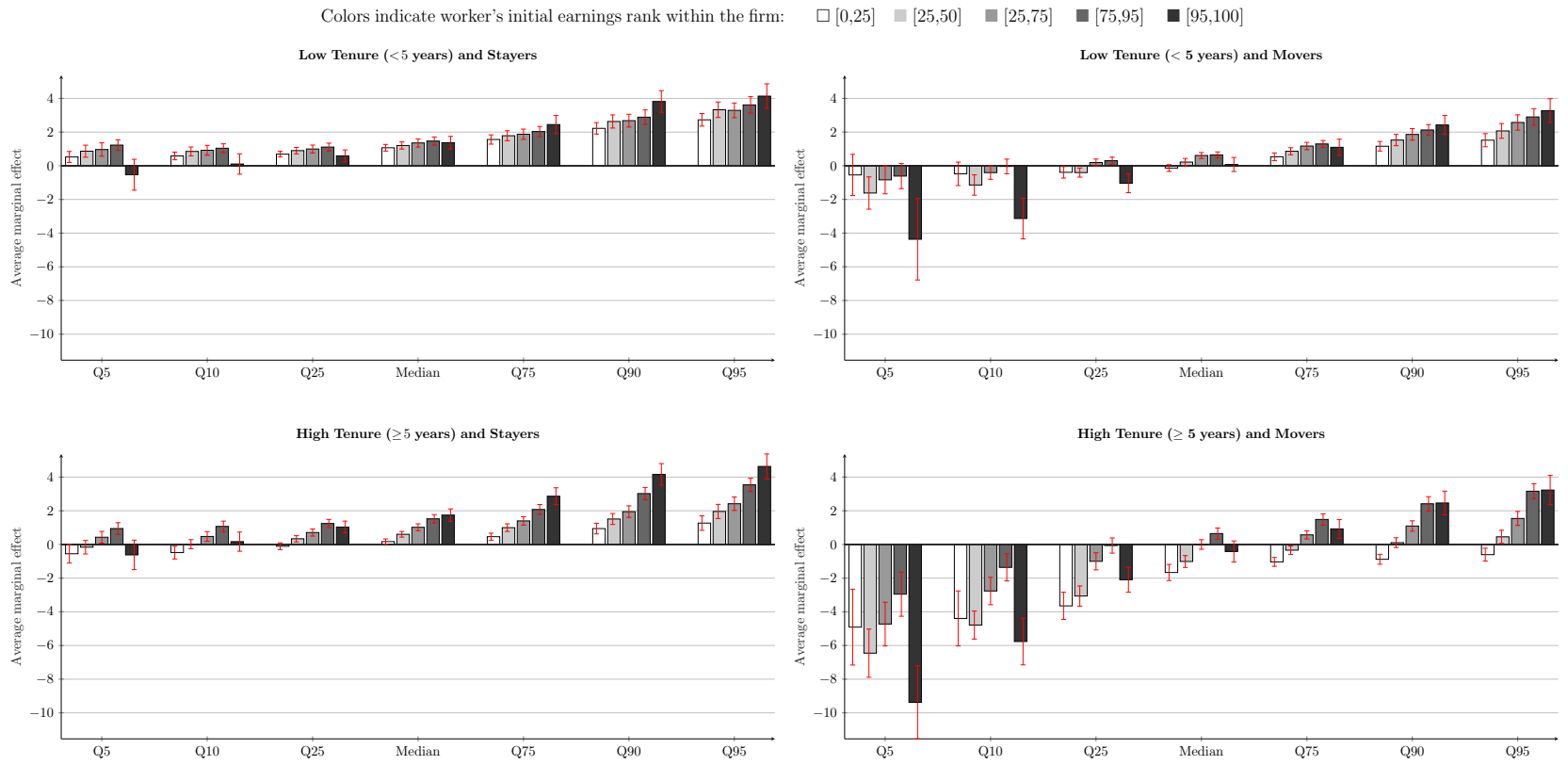
Note: Figure plots the average marginal effects of firm—and competitor—innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels. In addition to the own firm and competitor innovation measures from the baseline specification, we also include controls for own firm and competitor stock returns and also allow these coefficients to vary with a worker's rank within the firm. Estimates are standardized to correspond with one standard deviation effects for each variable of interest. The worker earnings rank is defined net of deterministic life cycle effects. We focus on 5-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.20: Earnings growth and competitor innovation conditional on worker earnings levels and tenure: movers versus continuing workers



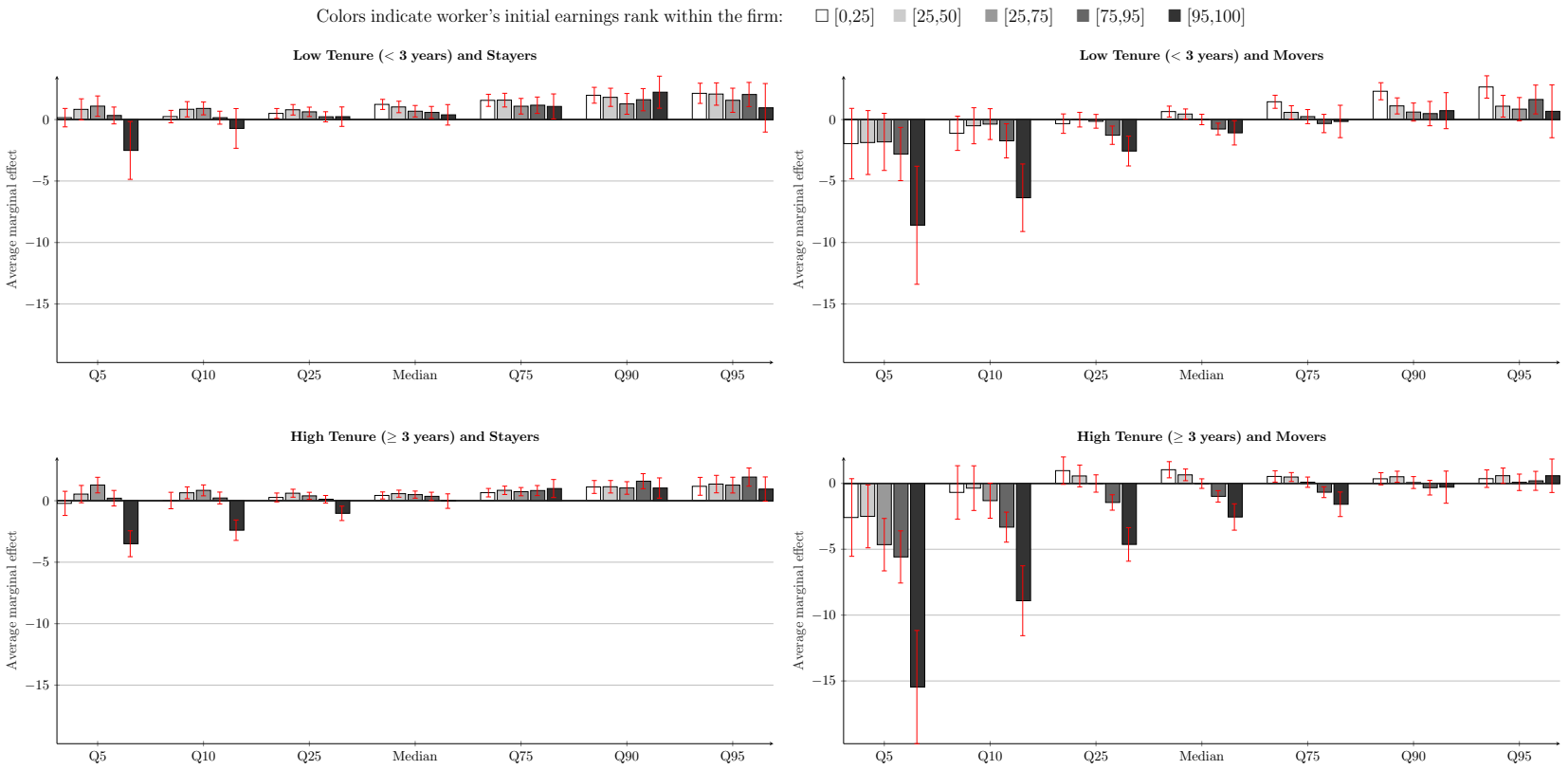
Note: Figure plots the average marginal effects of firm innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels and years of tenure with the firm. Workers are sorted into two groups based upon whether they have less than 3 years or greater than or equal to 3 years of tenure, and we allow for separate coefficients for each tenure × lagged earnings bin. Estimates are standardized to correspond with one standard deviation effects for 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 A.21: Earnings growth and own firm innovation conditional on worker earnings levels and tenure: movers versus continuing workers (5 year cutoff)



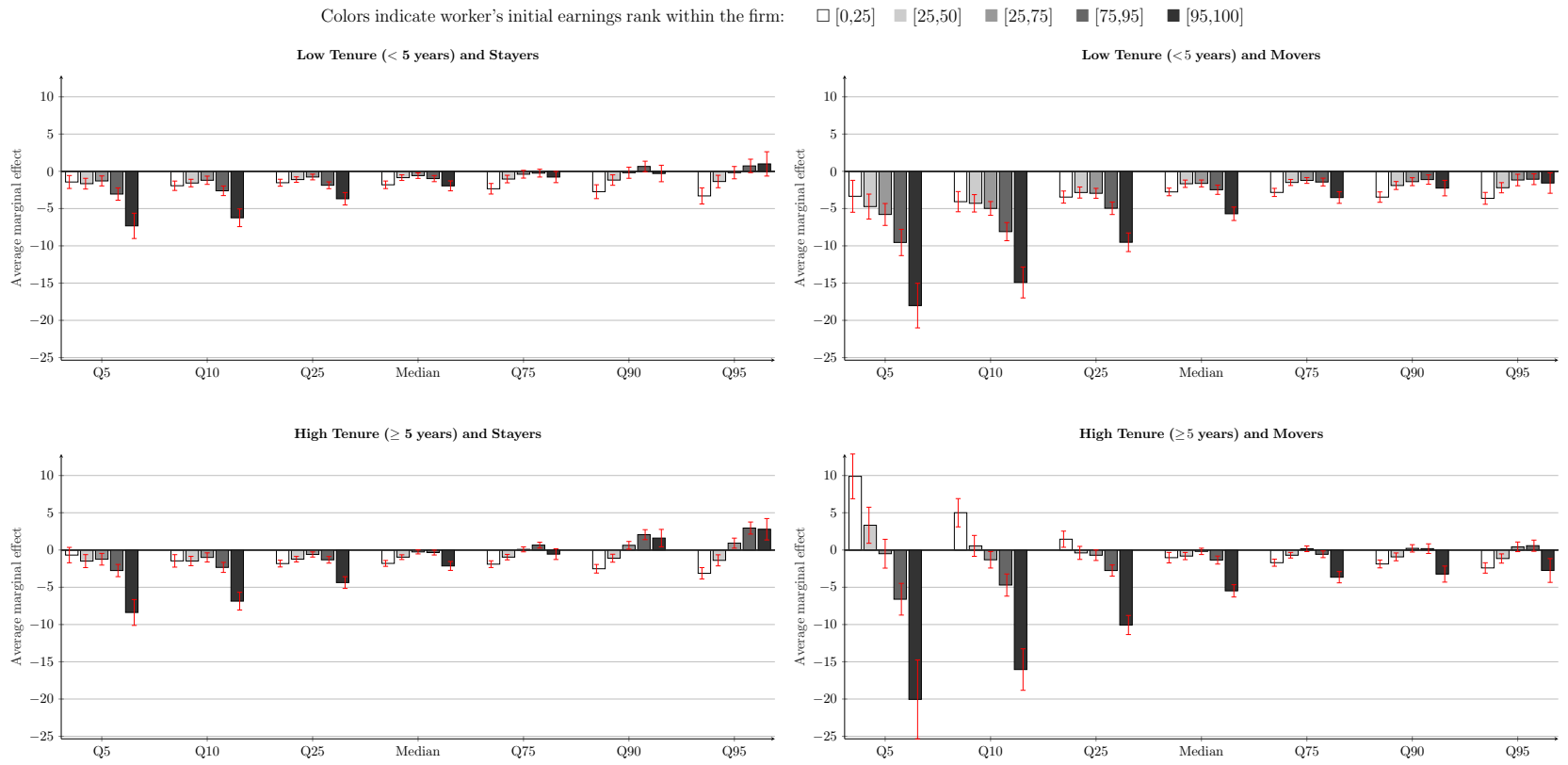
Note: Figure plots the average marginal effects of firm innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels and years of tenure with the firm. Workers are sorted into two groups based upon whether they have less than 5 years or greater than or equal to 5 years of tenure, and we allow for separate coefficients for each tenure × lagged earnings bin. Estimates are standardized to correspond with one standard deviation effects for 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-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.22: Earnings growth and own firm process innovation conditional on worker earnings levels and tenure: movers versus continuing workers



Note: Figure plots the average marginal effects of firm process innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels and for workers with different earnings levels and years of tenure with the firm. Workers are sorted into two groups based upon whether they have less than 3 years or greater than or equal to 3 years of tenure, and we allow for separate coefficients for each tenure × lagged earnings bin. Estimates are standardized to correspond with one standard deviation effects for 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-yr growth rates. The units on the vertical axis correspond to log points (times 100).

Figure A.23: Earnings growth and competitor innovation conditional on worker earnings levels and tenure: movers versus continuing workers (5 year cutoff)



Note: Figure plots the average marginal effects of firm innovation that are implied by the quantile regression estimates (equation (12) in the main text) for workers with different earnings levels and years of tenure with the firm. Workers are sorted into two groups based upon whether they have less than 5 years or greater than or equal to 5 years of tenure, and we allow for separate coefficients for each tenure × lagged earnings bin. Estimates are standardized to correspond with one standard deviation effects for 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).

Table A.1: Worker descriptive statistics: Full versus matched sample

Panel A. Matched sample

	Obs	Mean	SD	1%	5%	10%	25%	50%	75%	90%	95%	99%
Earnings (in thousands of 2013 dollars)	14,621,600	74.2	146.6	4.8	15.9	24.3	39.3	57.6	82.8	123.2	165.4	343.5
Age	14,621,600	39.6	8	26	27	29	33	39	46	51	53	54
Firm tenure	14,621,600	6.2	5.2	1	1	1	2	5	9	14	17	23
Firm tenure $\geq$ 3 years	14,621,600	0.7	0.4	0	0	0	0	1	1	1	1	1
Firm tenure $\geq$ 5 years	14,621,600	0.5	0.5	0	0	0	0	1	1	1	1	1
Cumulative 3 year log residual earnings growth	14,593,600	-0.07	0.59	-2.31	-0.88	-0.53	-0.17	-0.01	0.13	0.38	0.58	1.1
Cumulative 5 year log residual earnings growth	13,532,500	-0.09	0.61	-2.38	-0.96	-0.59	-0.21	-0.01	0.15	0.38	0.58	1.11
Cumulative 10 year log residual earnings growth	10,675,100	-0.12	0.65	-2.52	-1.09	-0.69	-0.28	-0.03	0.17	0.41	0.6	1.16
Left firm after 1 year	14,621,600	0.153	0.36	0	0	0	0	0	0	1	1	1
Left firm after 3 years	14,621,600	0.337	0.473	0	0	0	0	0	1	1	1	1
Number of years with zero earnings	13,823,100	0.142	0.566	0	0	0	0	0	0	0	1	3
Number of years with zero earnings, conditional on leaving firm after 3 years	4,661,500	0.327	0.875	0	0	0	0	0	0	1	3	4
Application for DI	11,128,600	0.026	0.159	0	0	0	0	0	0	0	0	1
Application for DI, conditional on leaving firm after 3 years	3,988,800	0.041	0.199	0	0	0	0	0	0	0	0	1

Panel B. SSA worker sample (based on 10% sample)

	Obs	Mean	SD	1%	5%	10%	25%	50%	75%	90%	95%	99%
Earnings (in thousands of 2013 dollars)	110927670	58.2	124.1	2.7	7.8	13.8	26.5	43.4	66	100.3	138.2	313.5
Age	104,030,810	38.9	8.1	26	27	28	32	38	46	51	52	54
Firm tenure	110,758,800	5.1	4.7	1	1	1	2	3	7	12	15	21
Firm tenure $\geq$ 3 years	110,758,800	0.6	0.5	0	0	0	0	1	1	1	1	1
Firm tenure $\geq$ 5 years	110,758,800	0.4	0.5	0	0	0	0	0	1	1	1	1
Cumulative 3 year log residual earnings growth	103,627,700	-0.09	0.65	-2.63	-1.09	-0.64	-0.21	-0.01	0.15	0.43	0.66	1.27
Cumulative 5 year log residual earnings growth	93,147,900	-0.1	0.67	-2.69	-1.15	-0.69	-0.25	-0.02	0.17	0.44	0.67	1.3
Cumulative 10 year log residual earnings growth	68,598,500	-0.12	0.71	-2.8	-1.26	-0.77	-0.3	-0.02	0.19	0.47	0.71	1.38
Left firm after 1 year	110,535,700	0.249	0.432	0	0	0	0	0	0	1	1	1
Left firm after 3 years	110,013,200	0.454	0.498	0	0	0	0	0	1	1	1	1
Number of years with zero earnings	101,607,000	0.248	0.792	0	0	0	0	0	0	1	2	4
Number of years with zero earnings, conditional on leaving firm after 3 years	46,355,400	0.467	1.085	0	0	0	0	0	0	2	3	5
Application for DI	88,363,000	0.031	0.173	0	0	0	0	0	0	0	0	1
Application for DI, conditional on leaving firm after 3 years	39,136,500	0.048	0.213	0	0	0	0	0	0	0	0	1

Note: Table reports univariate summary statistics for the sample of matched (Panel A) and unmatched (Panel B) worker-level measures. The unit of analysis is the worker-year. Sample sizes have been rounded to the nearest 100.

Table A.2: Firm descriptive statistics: matched vs non-matched sample

A. Matched sample												
	Obs	Mean	SD	1%	5%	10%	25%	50%	75%	90%	95%	99%
Employment (1000s)	117,300	8.42	36.88	0.01	0.02	0.05	0.17	0.79	4	16.5	37.9	128.9
Employment (SSA, 1000s)	119,900	3.77	16.34	0	0.01	0.02	0.08	0.34	1.7	7.0	16.0	59.2
Book assets, log	119,900	4.87	2.24	0.46	1.4	2.0	3.3	4.7	6.4	7.9	8.8	10.3
Return on assets	119,500	-0.01	0.28	-1.31	-0.55	-0.27	-0.01	0.07	0.12	0.17	0.2	0.29
R&D to assets	73,600	0.09	0.14	0	0	0	0.01	0.04	0.11	0.23	0.37	0.75
Firm volatility	110,600	-3.4	0.56	-4.54	-4.29	-4.13	-3.81	-3.41	-3.02	-2.65	-2.44	-2.05
Firm stock return	118,600	0.14	0.74	-0.86	-0.68	-0.54	-0.27	0.03	0.37	0.85	1.34	2.85
Industry stock return	118,300	0.13	0.29	-0.51	-0.33	-0.21	-0.05	0.12	0.29	0.47	0.62	0.9
Firm innovation	119,900	0.06	0.24	0	0	0	0	0	0.01	0.12	0.31	1.12
Firm innovation, non-process	119,900	0.03	0.13	0	0	0	0	0	0	0.06	0.17	0.64
Firm innovation, process	119,900	0.02	0.08	0	0	0	0	0	0	0.03	0.09	0.39
Competitor innovation	119,900	0.15	0.27	0	0	0	0	0.04	0.19	0.43	0.7	1.19
B. Non-Matched sample												
	Obs	Mean	SD	1%	5%	10%	25%	50%	75%	90%	95%	99%
Employment (1000s)	24,900	5.84	30.6	0	0	0.02	0.11	0.63	2.8	9.9	21.8	108.8
Book assets, log	25,900	4.89	2.19	0.4	1.4	2	3.3	4.9	6.4	7.7	8.6	10.2
Return on assets	25,700	-0.02	0.29	-1.31	-0.57	-0.28	-0.02	0.06	0.11	0.17	0.21	0.31
R&D to assets	12,200	0.08	0.14	0	0	0	0	0.02	0.09	0.22	0.35	0.76
Firm volatility	19,500	-3.28	0.56	-4.51	-4.18	-3.99	-3.68	-3.3	-2.9	-2.53	-2.33	-1.94
Firm stock return	25,500	0.1	0.71	-0.87	-0.72	-0.59	-0.31	0	0.33	0.82	1.33	2.67
Industry stock return	25,300	0.11	0.27	-0.51	-0.33	-0.21	-0.05	0.11	0.26	0.43	0.59	0.86
Firm innovation	25,900	0.01	0.11	0	0	0	0	0	0	0	0.03	0.34
Firm innovation, non-process	25,900	0.01	0.05	0	0	0	0	0	0	0	0.01	0.15
Firm innovation, process	25,900	0	0.04	0	0	0	0	0	0	0	0	0.09
Competitor innovation	25,900	0.1	0.2	0	0	0	0	0.01	0.13	0.29	0.43	1.07

Note: Table reports univariate summary statistics for the sample of matched (Panel A) and unmatched (Panel B) public firms. The unit of analysis is the GVKEY-year. Sample sizes have been rounded to the nearest 100.

Table A.3: 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 (3) 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 (4)). 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 A.4: Earnings growth and own/competitor innovation, stock returns, and profit growth conditional on earnings levels: OLS estimates

A. Innovation						
Variable	Scaling	Earnings rank				
		[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	StdDev	1.19	1.09	1.40	1.85	1.53
	Elasticity	0.149	0.137	0.176	0.231	0.191
		[10.86]	[8.07]	[9.98]	[14.41]	[8.34]
Innovation by competitors, $A_{I\setminus f}$	StdDev	-1.92	-1.40	-1.01	-2.20	-5.92
	Elasticity	0.39	0.284	0.204	0.444	1.201
		[-6.41]	[-5.68]	[-3.91]	[-6.87]	[-11.9]
B. Stock returns						
Variable	Scaling	Earnings rank				
		[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Own firm stock return, R_f	Elasticity	0.0355	0.0314	0.0306	0.0403	0.0673
		[15.21]	[15.91]	[13.66]	[14.07]	[11.85]
Competitor stock return, $R_{I\setminus f}$	Elasticity	0.0570	0.0236	0.0218	0.0238	0.0485
		[7.26]	[3.99]	[4.11]	[3.29]	[3.06]
C. Profit growth over the next 5 years						
Variable	Scaling	Earnings rank				
		[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Own firm profit growth, $\log\left[\frac{1}{\lvert 5\rvert}\sum_{\tau=1}^{5} X_{f,t+\tau}\right] - \log X_{f,t}$	Elasticity	0.0921	0.0706	0.0672	0.0728	0.1283
		[33.129]	[28.016]	[26.047]	[18.909]	[17.456]
Competitor profit growth $\log\left[\frac{1}{\lvert 5\rvert}\sum_{\tau=1}^{5} X_{I\setminus f,t+\tau}\right] - \log X_{f,t}$	Elasticity	0.0092	0.0049	0.0023	0.0002	0.0021
		[5.274]	[3.829]	[1.513]	[0.119]	[0.814]

Note: Table plots the average marginal effects of firm—and competitor—innovation that are implied by OLS estimates of equation (12) in the main text, where we allow for heterogenous coefficients for workers with different earnings levels. In panel A, we use our baseline own and competitor innovation measures, and estimates are scaled to correspond with a one standard deviation changes in each (analogous to Figure 3 in the main text). We additionally report the ratio of these coefficients to the firm-level slope coefficients from Table 1. In panels B and C, we report coefficients for analogous specifications where we use own firm and competitor stock returns and 5 year cumulative profit growth, respectively. 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).

Table A.5: Earnings growth and own/competitor innovation or stock returns conditional on earnings levels and mobility status: OLS estimates

A. Innovation							
Variable	Mobility Status	Scaling	Earnings rank				
			[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Innovation by the firm, A_f	Overall	StdDev	1.19	1.09	1.40	1.85	1.53
		Elasticity	0.149	0.137	0.176	0.231	0.191
			[10.86]	[8.07]	[9.98]	[14.41]	[8.34]
	Stayer	StdDev	1.08	1.38	1.61	2.08	2.16
		Elasticity	0.135	0.173	0.201	0.261	0.271
			[7.48]	[9.79]	[10.46]	[14.07]	[9.22]
	Switcher	StdDev	-0.28	-0.23	0.38	0.75	-1.04
		Elasticity	-0.035	-0.028	0.047	0.093	-0.13
			[-1.92]	[-1.5]	[3.11]	[4.97]	[-2.65]
Innovation by competitors, $A_{I\setminus f}$	Overall	StdDev	-1.92	-1.40	-1.01	-2.20	-5.92
		Elasticity	0.39	0.284	0.204	0.444	1.201
			[-6.41]	[-5.68]	[-3.91]	[-6.87]	[-11.9]
	Stayer	StdDev	-2.07	-1.37	-0.60	-1.22	-4.21
		Elasticity	0.419	0.279	0.122	0.248	0.854
			[-6.39]	[-6.12]	[-3.11]	[-5]	[-8.76]
	Switcher	StdDev	-1.90	-1.31	-1.25	-3.05	-7.99
		Elasticity	0.385	0.266	0.253	0.618	1.622
			[-6.82]	[-5.22]	[-5.06]	[-8.18]	[-13.18]
B. Stock returns							
Variable	Mobility Status	Scaling	Earnings rank				
			[0,25]	[25,50]	[50,75]	[75,95]	[95,100]
Own firm stock return, R_f	Overall	Elasticity	0.0355	0.0314	0.0306	0.0403	0.0673
			[15.21]	[15.91]	[13.66]	[14.07]	[11.85]
	Stayer	Elasticity	0.0328	0.0293	0.0323	0.0427	0.0686
			[13.17]	[13.49]	[12.05]	[14.48]	[13.1]
	Switcher	Elasticity	0.0113	0.0129	0.0106	0.018	0.042
			[2.94]	[4.65]	[4.41]	[5.77]	[6.19]
Competitor stock return, $R_{I\setminus f}$	Overall	Elasticity	0.0570	0.0236	0.0218	0.0238	0.0485
			[7.26]	[3.99]	[4.11]	[3.29]	[3.06]
	Stayer	Elasticity	0.0128	0.0088	0.0122	0.0193	0.0532
			[2.13]	[1.96]	[2.76]	[3.02]	[4.11]
	Switcher	Elasticity	0.0775	0.038	0.0351	0.0389	0.0509
			[9.03]	[4.6]	[4.35]	[3.94]	[1.99]

Note: Table plots the average marginal effects of firm—and competitor—innovation that are implied by OLS estimates of equation (12) in the main text, where we allow for heterogenous coefficients for workers with different earnings levels. We also report estimates where the equation is estimated separately for workers that remain with the firm (stayers) versus workers that leave the firm (switchers). In panel A, we use our baseline own and competitor innovation measures, and estimates are scaled to correspond with a one standard deviation changes in each (analogous to Figure 3 in the main text). We additionally report the ratio of these coefficients to the firm-level slope coefficients from Table 1 ('elasticity'). In panel B, we report coefficients for analogous specifications where we replace the innovation measures with own firm and competitor stock returns. 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).