In Search of Ideas: Technological Innovation and Executive Pay Inequality*

Carola Frydman† Dimitris Papanikolaou‡

Abstract

We develop a general equilibrium model that delivers realistic fluctuations in both the level and the dispersion in executive pay as a result of changes in the technology frontier. In our model, executives add value to the firm not only by participating in production decisions, as do other workers in the economy, but also by identifying new investment opportunities. The economic value of these two distinct components of the executives’ job varies with the state of the economy. Improvements in technology that are specific to new vintages of capital raise the return to managers’ skills for discovering new growth projects, and thus increase the compensation of executives relative to workers. When most of the dispersion in managerial skills lies in the ability to find new projects, disparities in pay across executives also increases in response to these embodied technological shocks. Our model delivers testable predictions about the relation between executive pay and growth opportunities that are quantitatively consistent with historical and modern data on executive pay.

*We thank Claudia Olivetti, Amit Seru, Andrei Shleifer, Per Stromberg, and the seminar participants at Boston College, FRB Chicago, Copenhagen Business School, University of Geneva, NYU Economics, and Stockholm School of Economics for helpful comments and discussions. We thank Kevin J. Murphy for providing the Forbes compensation data. Dimitris Papanikolaou thanks Amazon (AWS in Education Grant award) for research support.

†Kellogg School of Management and NBER, c-frydman@kellogg.northwestern.edu
‡Kellogg School of Management and NBER, d-papanikolaou@kellogg.northwestern.edu
The dispersion in pay between top executives and workers, as well as among executives, has fluctuated considerably over the last century. Over this period, pay inequality in America has followed a well-documented J-shaped pattern (Frydman and Saks, 2010). A less well-known fact is that the fluctuations in relative pay at medium-run frequencies are also striking. Understanding the underlying factors that drive these medium-run fluctuations can shed light into the forces behind the movements in overall pay inequality. We propose that technological innovation, and its impact on the value of investment opportunities in the economy, is an important driver of this process. We develop an equilibrium model of executive pay that links both the level and the dispersion in executive compensation to the current state of the economy. The key insight of our model is that executives contribute to their firms along multiple dimensions; importantly, the marginal value of managerial skills changes with the technology frontier, leading to substantial fluctuations in both the level and the dispersion in executive pay over time.

We build a dynamic general equilibrium model with heterogeneous firms that employ executives and workers. Executives add value to the firm along two dimensions. First, similar to production workers, they provide labor services that are complementary to the firms’ existing assets. Second, executives also participate in the creation of new capital by identifying new investment opportunities for the firm. The efficiency of an executive in identifying these opportunities depends on the quality of the match between the firm and the executive. Matching between executives and firms is random, and so the quality of the match is initially unobservable. Over time, as executives make investment decisions, all market participants update their beliefs about the quality of the match based on their observed performance, and managers with poor performance are fired. In equilibrium, executives are rewarded for both of their skills, while workers are only rewarded for their efforts in production. Similar to worker compensation, executive pay includes a component that is related to their direct contribution to the production process, which is proportional to aggregate output. But the compensation of executives includes a second component that depends on the marginal return to new investments, which in turn depends on the perceived quality of the match and the bargaining power of executives.

Our model generates significant time variation in both the level of executive pay – scaled by either the earnings of the average worker, or by total output – and in the dispersion in pay among executives. The key mechanism is that the marginal returns to these two skills are neither constant nor comove perfectly with each other. This result arises naturally in our model because our economy is characterized by two forms of technological progress. Some technical advances take the form of improvements in labor productivity, and are complementary to existing investments. This type of...

---

1These medium-run fluctuations account for approximately one-third of the overall fluctuations in pay dispersion between executives and workers, as well as between executives across different firms. Specifically, the time-series standard deviation of the pay differences between executives and workers (measured by the log ratio of mean executive pay to mean worker pay) and among executives (the log cross-sectional standard deviation of executive compensation across firms) over the 1936-2005 sample of Frydman and Saks (2010) is 50% and 28%, respectively. We filter those series using the band-pass filter and keep the components that correspond to medium-run frequencies – cycles between 5 and 50 years. The time-series volatility of these two components is 15% and 10%, respectively.
technological progress, which we refer to as disembodied technical change, benefits both workers and executives. Other types of innovations are embodied in new vintages of capital—we refer to this shock as embodied technical progress. This form of technical change leads to fluctuations in the marginal return of new investments that are contemporaneously uncorrelated with aggregate output. That is, these technological advances increase output only after they are implemented through the formation of new capital stock. Since executives take part in discovering new investment opportunities, their compensation reacts immediately to embodied technical progress, but the remuneration of workers only does so with delay. Thus, the level of pay of the average executive relative to the earnings of the average worker increases with the ratio of the marginal return to new investments relative to current output. Since the quality of the executive-firm match determines the managers’ ability to identify new growth prospects for their firms, the dispersion in pay across executives also comoves with the level of relative pay.

We estimate the parameters of the model using the simulated method of moments. We target moments of aggregate investment and consumption, and the dispersion in firm-level investment rates, valuations and profitability. In addition, we also target features of executive pay. Our model generates a realistic dispersion of executive pay across firms, and substantial time variation in both the level and dispersion of executive compensation. In terms of magnitudes, our model can replicate between one-third and one-half of the observed fluctuations in pay inequality. Our parameter estimates imply that both dimensions through which executives add value to their firms are important determinants of pay: on average, identifying new growth opportunities accounts for approximately 58% of executive pay.

Our model also delivers testable predictions about the relation between executive pay and firm growth in the cross-section of firms. In particular, our model implies that executive pay should be higher in fast growing firms. We examine these predictions by using two main datasets on executive pay. First, we use Execucomp, which provides information on executive compensation for a large number of publicly-traded firms since 1992. Second, we use the long-run dataset from 1936 to 2005 constructed by Frydman and Saks (2010). A main advantage of the historical data is that they allow us to study a much longer time period, and therefore provide more variation in aggregate conditions. However, these data cover a much smaller number of firms, and have limited industry variation. When possible, we present our analysis using both datasets.

We examine the relation between the level of executive pay and firm growth opportunities in several ways. First, we show that, controlling for firm size, an increase in executive pay predicts future firm growth. Second, we show that executive compensation is correlated with various measures of growth opportunities at the firm level, including investment, Tobin’s Q or the estimated value of

\[2\] With a fixed set of parameters over a 80-year period, our model cannot fully explain the low-frequency component in pay disparities over the twentieth century. Frydman and Saks (2010) show that executive pay inequality exhibits a J-shaped pattern over this period, with a sharp decline in the 1940s, a period of little dispersion in pay from the 1950s to the 1970s, and a rapid increase in inequality since the 1980s. It is quite likely that changing some of the parameters over time (for example, by reducing the executives’ bargaining power during the 1940s and increasing it since the 1980s), our model could better match the low-frequency dynamics of income inequality.
new innovations of Kogan, Papanikolaou, Seru, and Stoffman (2012).\textsuperscript{3} Third, we exploit changes in tax policy that led to variation in the effective investment tax credit (ITC) at the industry level as a source of plausibly exogenous variation in the value of investment opportunities available to firms. Overall, we find a statistically significant and economically substantial relationship between the level of executive pay and firm growth opportunities, even after we control for a variety of firm observable characteristics, including firm size and current profitability (as well year and industry or firm fixed effects). To evaluate the quantitative plausibility of our proposed mechanism, we replicate our key empirical results in simulated data from the model. We find that the magnitude of the estimated correlations is quantitatively similar between the model and the data.

In addition to these cross-sectional predictions, our model also has sharp predictions about the aggregate dynamics of executive pay inequality across firms, as well as between executives and workers. Specifically, our model implies that both the level and the dispersion in executive pay are positively related to the return to new investments scaled by total output. Even though this ratio is not directly observable in the data, in our model it is positively related to several variables that are instead observable, as the average rate of investment in the economy. Thus, we construct a model-based proxy for the aggregate level of investment opportunities in the economy, and show that this proxy has a strong association with the level and dispersion in executive pay at medium-run frequencies (frequencies of 5 to 50 years), with correlations in excess of 75 percent. This fact is consistent with the idea that shocks to the value of investment opportunities are the main driver behind these medium-run movements in pay inequality, which themselves account for approximately one-third of the total fluctuations in executive pay inequality from the 1930s to the end of our sample period in the early 2000s.

Finally, we also examine the relationship between executive pay inequality and various measures of investment opportunities across different industries. This exercise allows us to ‘difference’ out any slow-moving, common factors (such as shifts in corporate governance) that may play a role in driving the aggregate trends in pay inequality. Consistent with our model, we find that pay inequality – again, both between executives and workers, as well as among executives across firms – is positively related to measures of investment opportunities at the industry level, such as Tobin’s $Q$, investment, and the the economic value of firms’ innovative activity proposed by Kogan et al. (2012). Our coefficient estimates imply that fluctuations in investment opportunities at the industry level have a similar impact on the level and dispersion of executive pay as an increase in average firm size.

In sum, an important contribution of our study is to show that a relatively stripped-down model of technological innovation and growth can go a long way towards quantitatively replicating the aggregate and cross-sectional dynamics of executive pay inequality. Our framework contains no

\textsuperscript{3}The correlation between the level of executive pay and Tobin’s $Q$ is well-documented in the empirical literature on executive pay (see Smith and Watts (1992) for American firms and Fernandes, Ferreira, Matos, and Murphy (2013) for international evidence). However, this literature does not provide a theoretical foundation for these documented correlations.
structural shifts in parameters, and therefore the variation over time arises purely through the stochastic nature of the model. In practice, such structural shifts may have affected the bargaining power between shareholders and executives, and including them in the model may help better accommodate the long run trends in executive pay. Examples of factors that may have caused structural shifts over the century include: changes in the taxation of top incomes (Frydman and Molloy, 2011), the power of labor unions (Frydman and Molloy, 2012), corporate governance (Kaplan, 2013), the supply of talent (Goldin and Katz, 2008), and the portability of managerial skills across firms (Murphy and Zábojník, 2004; Frydman, 2015).

The long-run trends in the disparity of pay between top executives and workers, as well as among executives, have sparked a considerable debate among economists and policymakers. Proponents of the ‘rent-extraction’ view of compensation propose that executive pay is the result of weak corporate governance that allows managers to extract excessive compensation relative to the value that they add to their firms. By contrast, advocates of a ‘market-based’ view of executive compensation argue that the observed levels of pay are the efficient outcome from firms competing for scarce managerial talent in the market for executives. Our work contributes to this debate by proposing a new market-based model of equilibrium pay that quantitatively accounts for a substantial fraction of the fluctuations in pay. Our framework also delivers testable predictions relating the level and dispersion of executive pay to the level of investment opportunities in firms or sectors. While it is difficult to rule out that these particular relationships are not the result of managerial rent-seeking behavior, the facts that we document seem inconsistent with ‘naive’ versions of the rent extraction hypothesis—for instance, the case in which managers extract a constant fraction of firm value.

Recent theoretical studies on the determinants of executive pay emphasize the role of competitive assignment models, which propose that managerial innate ability is complementary to firm size (Terviö, 2008; Gabaix and Landier, 2008). In a competitive labor market for executives, large firms are willing to remunerate executives handsomely to attract the most talented individuals; even small differences in managerial ability can lead to substantial differences in pay across firms. These static models predict a positive association between the compensation of executives and the current size of their firms. Our dynamic framework delivers similar predictions for the relationship between executive pay and a ‘long-run’ notion of firm size that encompasses not only the value of assets in place but also the firm’s future growth due to new investments. Thus, we relate executive pay not only to the current size of the firm, but also to the growth in firm size. An important contribution of our study is therefore to help explain some of the variation in executive compensation that remains

---

4For instance, starting in August 2015, the Securities and Exchange Commission (SEC) requires all public firms to disclose the ratio of the pay of the CEO to the median compensation of their employees.

5Our setting abstracts from assortative matching considerations because managerial skills are match-specific. We make this assumption primarily to simplify the computation of the equilibrium. Alternatively, we could allow for heterogeneity in growth opportunities across firms, and for heterogeneity in managers’ innate ability to find new projects. While this model would deliver assortative matching, solving its dynamic version would be computationally challenging: following any change in the relative ranking of firms in terms of investment opportunities, all executives would optimally switch firms.
across firms even after accounting for the current size of the executives’ firms. In our setting, the dispersion in executive pay is driven by the managers’ skills at finding new growth opportunities for their current firm – and these skills are more valuable at times of greater expansion in the technology frontier. Since firms that grow faster will eventually be larger, our model provides another foundation for the relationship between pay and size beyond the assumption that managerial skill and firm size are complements in the production process.

Our model combines features from several strands of the literature. We differ from most standard models of executive pay in that we consider the case of executives possessing more than one skill, and we allow for the prices of these skills to vary over time. Conceptually, the paper closest to ours is Eisfeldt and Kuhnen (2013). Relative to their work, however, we consider specific managerial skills (that is, working in production and identifying new growth opportunities), which allows us to take the model directly to the data. Moreover, relative to their work, the prices for these skills arise endogenously in our equilibrium model, and we focus on pay inequality. Our work is also related to Lustig, Syverson, and Van Nieuwerburgh (2011), who also present a general equilibrium model of executive compensation with both embodied and general-purpose technological progress. In their model, executives are paid proportionally to the value of installed capital. Lustig et al. (2011) show that a shift in the composition of productivity growth from capital-embodied to general-purpose leads to an increase in managerial pay. By contrast, in our setting, the level of executive compensation depends also on the value of new growth opportunities, which leads to different testable implications. Finally, our work is methodologically related to Taylor (2010, 2013), who estimates a structural model of executive wages and turnover. While he focuses on reduced-form relations that can arise in partial equilibrium models, we propose and estimate a general equilibrium model that has specific predictions about the relation between executive pay and firm growth. Despite considering very different modeling frameworks, we reach similar conclusions regarding the estimated parameters for the fraction of the match surplus that is captured by executives (about 50 percent). Moreover, we also find that large firing costs are needed in order to reproduce the level of turnover in the data (as in Taylor, 2010).

More broadly, our paper also contributes to the extensive literature studying the determinants of income inequality. Bakija, Cole, and Heim (2008) document that executives, managers and supervisors in the non-finance sector account for approximately 40% of those in the top 0.1% in terms of income in the 1979-2005 period, and approximately one-third of the top 1%. This group includes not only executives in large public firms, but also those working in closely-held businesses (C-corporations). Further, Kaplan and Rauh (2013) find that the wealthiest individuals in the US are increasingly composed of technology entrepreneurs. Even though our model is cast in terms of executives working for firms, it could as easily apply to entrepreneurs, since identifying new investment opportunities is arguably one of their defining characteristics. Conceptually, our model bears some similarity to models with skill-biased technical change (Griliches, 1969; Autor, Katz, and Krueger, 1998; Krusell, Ohanian, Ros-Rull, and Violante, 2000; Hornstein, Krusell, and
Violante, 2005). These studies argue that improvements in technology increase income inequality because skilled labor is more complementary to capital than unskilled labor. Similarly, in our setting managers’ ability to identify new valuable projects is complementary to technological progress embodied in new capital goods, whereas workers’ skills are not. We focus on a particular type of skilled labor (executives), that accounts for a large fraction of the top of the income distribution, and a particular skill (their ability to identify new projects), that is complementary to technological progress. We provide evidence consistent with our proposed mechanism by relating the level of executive pay to firm growth. In this sense, our work is related to the literature studying the links between creative destruction and income inequality (Jones and Kim, 2014; Kogan, Papanikolaou, and Stoffman, 2015; Aghion, Acgigit, Bergeaud, Blundell, and Hemous, 2015).

1 The Model

We consider a dynamic continuous-time economy. There is a continuum of firms of measure one, and time is indexed by $t$. We introduce households and firms in Sections 1.1 and 1.2, respectively. We discuss the role of executives in Section 1.3. Finally, we describe the competitive equilibrium of the model in Section 1.5.

1.1 Households and financial markets

The household side of the model is fairly standard. There is a continuum of households of measure $H > 1$. At any point in time, a subset of the households is employed as executives. For simplicity, we assume that each firm has one executive. There is a unit measure of firms; thus, the set of executives is also measure one. The subset of households that are not employed as executives inelastically supplies a homogenous flow of labor services equal to $h dt$. We refer to these $(H - 1)$ households as workers.

Executives manage the firms in our economy. Importantly, we propose that executives play two distinctive roles. First, they discover new investment opportunities and undertake investment decisions. Second, they operate the firms’ assets in place—for example, they manage the workers and installed capital to produce with the existing projects. To model the executives’ contribution to operating the firm’s existing assets, we assume that they are endowed with an effective flow of labor services equal to $e dt$. This activity captures all tasks associated with making efficient production decisions, and includes, but is not limited to, monitoring and managing workers, minimizing costs, or building an organizational structure that efficiently utilizes the existing workers. We normalize the total flow of labor services to one, and therefore $(H - 1) h + e = 1$. The parameters $e$ and $h$ help the model match the mean level of inequality between executives and workers. However, since

---

\textsuperscript{6}Though we describe the model as if one executive was allocated to each firm for simplicity, more generally we think of these executives as representing the team of top managers that makes investment decisions.
all managers provide the same level of effective labor services \( \epsilon \), this component of pay does not contribute to inequality among executives.

Households make consumption and savings decisions to optimize their lifetime utility of consumption. All households have the same preferences over sequences of consumption \( C \), given by

\[
J_t = E_t \left[ \int_t^\infty \log(C_s) \, ds \right].
\]

Households are not subject to liquidity constraints. They can sell their future labor income streams and invest the proceeds in financial claims.

Households have access to complete financial markets. Specifically, they can trade a complete set of state-contingent claims. We denote the equilibrium stochastic discount factor by \( \Lambda_t \), so the time-\( t \) market value of a time-\( T \) cash flow \( X_T \) is given by

\[
E_t \left[ \frac{\Lambda_T}{\Lambda_t} X_T \right].
\]

By considering the case of complete markets, we can focus solely on the behavior of a representative household which consumes the aggregate flow consumption \( C \) each period.

1.2 Firms, technology and aggregate output

There is a continuum of infinitely lived firms in the economy, which we index by \( f \in [0,1] \). Firms own and manage a collection of projects. Each project is the basic production unit in our economy. Each firm hires labor services—workers and executives—to operate the existing projects. The output of these projects can be used to produce either consumption or investment. New projects are created by combining investment goods (i.e., physical capital) and new ideas (i.e., investment opportunities). Investment goods are produced by firms, while ideas originate in executives.

Active projects

Each firm \( f \) owns a constantly evolving portfolio of projects, which we denote by \( \mathcal{J}_{ft} \). Projects are differentiated from each other by three characteristics: a) their operating scale, determined by the amount of capital goods associated with the project, \( k \); b) the systematic component of project productivity, \( \xi \); and c) the idiosyncratic, or project-specific, component of productivity, \( u \). Project \( j \), created at time \( \tau(j) \), produces a flow of output equal to

\[
y_{j,t} = \left( u_{j,t} e^{\xi_{\tau(j)}(j) k_{j,t}} \right)^\phi \left( e^{\xi_{1} L_{j,t}} \right)^{1 - \phi},
\]

where \( L_{j,t} \) is amount of labor allocated to this project. As we discuss in more detail below, the scale decision is made at the time of the project creation and it is irreversible. In contrast, the choice of labor \( L_{j,t} \) allocated to each project \( j \) can be freely adjusted every period. Firms purchase labor
services at the equilibrium wage \( w_t \). We denote by

\[
\pi_{j,t} = \sup_{L_{j,t}} \left[ \left( u_{j,t} e^{\xi_t(j)} k_{j,t} \right)^\phi \left( e^{x_t} L_{j,t} \right)^{1-\phi} - w_t L_{j,t} \right]
\]

the profit flow of project \( j \) under the optimal hiring policy.

We model technological progress as having heterogeneous effects on different vintages of capital. Specifically, technological innovations are characterized by two independent processes, \( \xi_t \) and \( x_t \). The shock \( \xi \) reflects technological progress embodied in new projects—that is, this shock does not affect the productivity of assets in place created with older technologies. We model \( \xi \) as an arithmetic random walk

\[
d\xi_t = \mu_\xi dt + \sigma_\xi dB_{\xi,t},
\]

where \( B_{\xi} \) is a standard Brownian motion independent of other shocks in the model. Note that \( \xi_s \) denotes the level of frontier technology at time \( s \). Thus, growth in \( \xi \) affects only the output of new projects created using the latest technological frontier. In this respect our model follows the standard vintage-capital model (Solow, 1960).

The second technology shock \( x_t \) is a standard labor-augmenting productivity process. Since labor is complementary to capital, \( x \) affects the output of all vintages of existing capital regardless of how distant they are to the technological frontier. The shock \( x \) also follows an arithmetic random walk

\[
dx_t = \mu_x dt + \sigma_x dB_{x,t}.
\]

where \( B_x \) is a standard Brownian motion independent of all other shocks in the model.

We model the productivity of each project as \( u_j \)—a stationary mean-reverting process that evolves according to

\[
du_{j,t} = \kappa_u (1 - u_{j,t}) dt + \sigma_u u_{j,t} dB_{u,t}^u,
\]

where \( B_{u}^u \) is a standard Brownian motion process independent of \( B_{\xi} \). We assume that \( dB_{j,t}^u dB_{j',t}^u = dt \) if projects \( j \) and \( j' \) belong in the same firm \( f \), and zero otherwise.\(^7\) All new projects implemented at time \( s \) start at the long-run average level of idiosyncratic productivity, i.e., \( u_{j,\tau(j)} = 1 \). Thus, all projects created at a point in time are ex-ante identical in terms of productivity, but differ ex-post due to the project-specific shocks.

The firm chooses the initial operating scale \( k \) of a new project irreversibly at the time of its creation. Firms cannot liquidate existing projects and recover their investment costs, but projects depreciate over time. Specifically, the scale of the project diminishes according to

\[
dk_{j,t} = -\delta k_{j,t} dt,
\]

\(^7\)We assume that the project productivity shocks are perfectly correlated at the firm level to ensure that the firm state vector is low-dimensional.
where $\delta$ is the economy-wide depreciation rate. Note that the aggregate (quality-adjusted) stock of installed capital in the economy $K$,

$$
K_t = \int_0^1 \left( \sum_{j \in J_{f,t}} e^{\xi_s(j)} k_{j,t} \right) df
$$

(9)

also depreciates at rate $\delta$.

Creation of new projects

To create a new project, firms must combine an investment opportunity with new investment goods. Executives identify investment opportunities and undertake the investment decisions. The frequency at which executives find new investment opportunities is match specific—that is, it depends on the quality of the match between the executive and her firm. Specifically, the likelihood of acquiring a new project is driven by a firm-specific Poisson process $N_{f,t}$ with arrival rate equal to $\lambda_{f,t}$. Depending on the quality of the match between the firm and its current executive, the arrival rate can be either high or low $\{\lambda_H, \lambda_L\}$, where $\lambda_H > \lambda_L$. Importantly, the arrival rate is unobservable to the firm, the executive, and all market participants, and there is no private information about the quality of the match. All parties learn about $\lambda_{f,t}$ by observing the firm’s investment decisions. In the next section, we describe the process through which firms and executives match and separate in more detail.

Once the firm acquires a new investment opportunity, it purchases new capital goods in quantity $I_{j,t}$ to implement a new project $j$ at time $t$. Investment in new projects is subject to decreasing returns to scale,

$$
k_{j,t} = I_{j,t}^\alpha.
$$

(10)

where $\alpha \in (0, 1)$ implies that investment costs are convex at the project level. Decreasing returns to investment imply that projects generate positive profits. We denote by

$$
q_t \equiv \sup_{k_{j,t}} \left\{ \mathbb{E}_t \left[ \int_t^\infty \frac{\Lambda_s}{\Lambda_t} \pi_{j,s} \, ds \right] - k_{j,t}^{1/\alpha} \right\}
$$

(11)

the net value of a new project implemented at time $t$ under the optimal investment policy, where $\Lambda_t$ is the equilibrium stochastic discount factor defined in Section 1.1. Since all projects created at time $t$ are ex-ante identical, $q$ is independent of $j$. Equation (11) also describes the value of a new investment opportunity that arrives at time $t$.

Aggregate output

The total output in the economy is equal to the aggregate of the output of all active projects,

$$
Y_t = \int_0^1 \left( \sum_{j \in J_{f,t}} y_{j,t} \right) df.
$$

(12)
Aggregate output can be allocated to either investment $I_t$ or consumption $C_t$,

$$Y_t = I_t + C_t. \quad (13)$$

The new investment goods $I_t$ produced at a point in time are used as inputs for the implementation of new projects, as given by the investment cost function defined in (10).

### 1.3 Executives

Executives participate in production decisions and identify new investment opportunities for the firm. The quality of a specific match determines the manager’s ability to discover new investment opportunities for her firm. Specifically, an executive is more likely to find a new idea if she is in a high-quality match than if the match is of poor quality. The firm and the executive learn about the quality of their match, although imperfectly, by observing the executive’s investment decisions, and firms fire the executives that perform poorly, as in Jovanovic (1979). The remuneration of an executive depends on the quality of the match. Since learning is imperfect, the quality of matches and the level executive pay vary across firms. Next, we describe the forces that determine the level of executive pay in more detail.

**Match Quality**

Firms employ up to one executive at a given point in time. In addition to providing a flow of labor services $e$ to operate the firm’s assets in place, the executive is in charge of discovering new investment opportunities. Recall that $\lambda_{f,t}$ determines the likelihood that the firm acquires a new investment opportunity at time $t$. We denote by $\lambda_{f,t} \in \{\lambda_L, \lambda_H\}$ the quality of the match between an executive and a firm. If a firm were to operate without an executive, it would obtain new investment opportunities at the expected rate $\lambda_L \, dt$. The quality of the match is firm-specific, and unobservable to all participants. We denote by $p_{f,t}$ the probability that the current match between the firm and the executive is of high quality—we refer to this measure as the **perceived quality** of the match.

**Learning and Executive Turnover**

After the executive is hired, the firm and the executive (as well as other market participants) learn about the quality of the match $p_{f,t}$ by observing the executive’s investment decisions. Standard results on filtering for point processes (Liptser and Shiryaev, 2001) imply that the evolution of $p_{f,t}$ is given by

$$dp_{f,t} = -p_{f,t} \left(1 - p_{f,t}\right) \lambda_D \, dt + \left(\frac{p_{f,t} \lambda_H}{\lambda_L + p_{f,t} \lambda_D} - p_{f,t}\right) dN_{f,t} \quad (14)$$

where $\lambda_D \equiv \lambda_H - \lambda_L$ is the difference in quality between a good and a bad match. Equation (14) shows that the perceived quality of the match $p_{f,t}$ increases sharply when the firm invests ($dN_{f,t} = 1$),
and drifts down slowly if the firm does not. Uncertainty about the quality of a given match is greatest for intermediate values of $p$. For these intermediate values, the firm’s beliefs are more likely to experience a greater change as new information comes along. For example, if the executive does not invest, the downward revision in beliefs will be largest when $p_{f,t}$ is close to 1/2.

The value of a match of perceived quality $p_{f,t}$ is

$$m_t(p_{f,t}) = p_{f,t} E_t \int_t^\tau \lambda_D \frac{\Lambda_s}{\Lambda_t} q_s ds$$

(15)

where $\tau$ is the (stochastic) time at which the match is dissolved, and $q_t$ is the net present value of a new project created at time $t$, which is defined in equation (11). Equation (17) describes the difference in firm value resulting from the firm hiring a new executive relative to the firm operating without any executive – in which case it finds projects at a rate $\lambda_t dt$.

Firms with poor-quality matches can choose to terminate the executives and replace them with someone new. Specifically, at any point in time, firms will fire the executives whose perceived match quality falls below a threshold $p_{f,t} \leq p^*_{f,t}$. Executives who are let go are thrown back in the pool of potential executives. Since the quality of the match is firm-specific, and it is not an innate characteristic of the executive (such as managerial ability), these managers can be potentially re-employed by a different firm. Since there is a continuum of firms and potential executives, the likelihood that the firm hires the same executive more than once is zero. The firm-specificity of match quality (which we assume is independent of the quality of the match between the same executive and a different firm) greatly simplifies solving the model. This assumption ensures that firms have no incentives to poach executives from other firms, and would likely hire executives that have been fired by other firms.

In addition to endogenous turnover, we also allow for exogenous separations: with flow probability $\beta$ each period, the match between CEO and firm is dissolved and the firm must hire – and train – a new executive. This assumption ensures that the distribution of match quality across executive-firm matches is stationary. In the absence of exogenous separations, firms would keep the first executive they ever hire who is a good match forever. The economy would quickly converge to an equilibrium where all matches are of high quality, and there would be no heterogeneity in match quality across firms.

Hiring Decisions

Since the quality of the match is firm-specific, all potential new executives are ex-ante identical. The firm has a prior belief that the quality of the match is high equal to $\bar{p}$. Training a new executive incurs a cost equal to $cm_t(\bar{p})$; that is, the training cost is proportional to the equilibrium value of a new match. For simplicity, we assume that this cost is a direct transfer from the firm to the households. This assumption guarantees that these costs do not affect the pool of aggregate resources available for consumption or investment.
Finally, we denote by $\lambda_t$ the mean quality among the current firm-executive matches,

$$\lambda_t \equiv \lambda_L + \lambda_D \int_0^1 p_{f,t} df.$$  \hfill (16)

Here, $\lambda_t$ affects the rate at which the economy acquires new projects – or equivalently the rate of capital accumulation. In general, the equilibrium level of $\lambda$ will depend on the efficiency of the firm-executive matching market – the mean of the distribution of $p_{f,t}$, or equivalently, the fraction of active matches that are of high quality.

**Executive Compensation**

Firms and executives bargain over the surplus generated by the match,

$$S_{f,t} \equiv m_t(p_{f,t}) - (1 - c)m_t(\bar{p}).$$  \hfill (17)

We assume that executives can capture a fraction $\eta$ of that surplus. Importantly, we make the simplifying assumption that the outside option of the executive is zero.\(^8\) That is, in order for the executives to agree to remain with the firm, the firm must promise to pay the executive a flow compensation $w_{f,t} dt$ – in addition to the compensation for labor services – that satisfies, for all $t$,

$$W_{f,t} = E_t \int_t^\tau \Lambda_s \Lambda_t w_{f,s} ds$$  \hfill (18)

and

$$W_{f,t} = \eta S_{f,t}.$$  \hfill (19)

As a result, the total compensation of an executive that works in firm $f$ is equal to

$$X_{f,t} = (e w_t + w_{f,t}) dt.$$  \hfill (20)

That is, executives are compensated for their effective labor services, at a price $w_t$, as well as for their ability to generate new investment opportunities at their current firm.

### 1.4 Discussion of the model’s assumptions

Before we proceed to the analysis of the equilibrium, we discuss some of the assumptions in our model. The key feature of our model is that executives add value by discovering new investment opportunities. The key feature of our model is that executives add value by discovering new investment opportunities in addition to participating in production decisions. Importantly, the marginal value of each of these two activities – as well as the ratio of their marginal values – varies

---

\(^8\)This assumption is made purely for analytical convenience. The assumption is equivalent to assuming that the measure of potential executives $H$ is sufficiently larger than the set of firms (whose measure is normalized to one) so that the discounted payoff for an unemployed executive from searching for a job to be zero, since the likelihood of being hired by a different firm in the near future is rather low. Alternatively, we could also have assumed that searching for a job entails a utility cost that is such that the continuation value of unemployment is normalized to zero.
over time. To obtain a meaningful distinction between the returns to new investments and the profitability of existing assets, as well as between the level of investment opportunities and the current size of firms, we use a model with vintage capital based on Kogan et al. (2015). This distinction between investment opportunities and current productivity is important because it allows us to show that the dispersion in executive pay depends not only on the current size (or profitability) of the firm, but also on the growth opportunities that the firm has. By contrast, in the neoclassical growth model, the returns to new investments are intimately related to the current size of the firm and the profitability of installed capital.

In addition to this feature, we have made several auxiliary assumptions to keep the model tractable. First, we abstract away from executive incentives. This choice is driven by our focus on disparities in the level, as opposed to the composition, of pay.

Second, we assume that projects arrive independently of the firms’ past investment decisions, and that firms incur convex investment costs at the project level. These two assumptions ensure that the optimal investment decision can be formulated as a static problem, and therefore that the cross-sectional distribution of firm size does not affect equilibrium aggregate quantities and prices. In the aggregate, the cost of investment is then a convex function of investment level, as in Abel (1983).

Third, we assume that conditional on their vintage, the quality of projects does not vary across firms. This assumption is made purely to keep the number of parameters manageable. Instead, we could allow for an idiosyncratic part to $\xi$ to capture the notion that the distribution of profitability of new ideas can be highly asymmetric. Our conjecture is that such an extension would lead to additional skewness in executive pay, without changing our predictions while introducing additional parameters to be estimated.

Fourth, our assumption that termination costs are proportional to the value of a new match implies that the firing threshold will not depend on the state of the economy, and it guarantees that the average match quality will be constant, thus reducing the aggregate state space and greatly simplifying our analysis. Relaxing this assumption would imply that the rate of turnover would vary with the state of the economy; the exact relation between turnover and economic growth would depend on the assumed functional form for the termination cost.

Last, we assume that the executive’s outside option is zero to keep the measure of firms constant over time. This assumption simplifies our analysis. Alternatively, we could extend the model by allowing the executive to leave the firm and start a new corporation. Since the value of a new firm will be proportional to $v_t(\bar{p})$, the qualitative predictions of this model would be similar, at the cost of having an expanding measure of firms.
1.5 Competitive equilibrium

Next, we describe the competitive equilibrium of our model. Our equilibrium definition is standard, and is summarized below.

**Definition 1** (Competitive Equilibrium). The competitive equilibrium is a sequence of quantities \( \{ C_t, I_t, Y_t, K_t \} \); prices \( \{ \Lambda_t, w_t \} \); household consumption decisions \( \{ C_{i,t} \} \); and firm investment and hiring decisions \( \{ I_{f,t}, L_{j,t}, p^* \} \) such that given the sequence of stochastic shocks \( \{ x_t, \xi_t, u_{j,t}, N_{f,t} \}, j \in \bigcup_{f \in [0,1]} J_{f,t}, f \in [0,1] \)  i) households choose consumption and savings plans to maximize their utility (1); ii) household budget constraints are satisfied; iii) firms maximize profits; iv) firms and investors rationally update match quality given (14); v) the executive’s continuation value satisfies (19), while flow executive pay \( \bar{w}_{f,t} \) satisfies the promise-keeping constraint (18); vi) the labor market clears, \( \int_0^1 \left( \sum_{j \in J_{f,t}} L_{j,t} \right) df = 1 \); vii) the demand for new investment equals supply, \( \int_0^1 I_{f,t} df = I_t \); viii) the market for consumption clears, and ix) the aggregate resource constraint (12) is satisfied.

We next characterize the equilibrium dynamics. For ease of exposition, we delegate all proofs to Appendix A.

We begin by showing that the firm’s termination threshold \( p^* \) is constant across firms and time. This result follows directly from our assumption of a proportional hiring/termination cost and greatly simplifies our analysis. Specifically, a firm will fire its executive if the value of the current match falls below the value of a new match, excluding training costs,

\[
p_{ft} \leq p^* \equiv (1 - c) \bar{p},
\]

or equivalently, when the match-specific surplus (17), which can be written as:

\[
S_{f,t} = (p_{ft} - p^*) \lambda_D E_t \int_t^T \frac{A_s}{K_t} q_s ds,
\]

becomes zero.

The fact that the firing threshold \( p^* \) is constant implies that the distribution of \( p_{ft} \) across firms is stationary and therefore the mean quality of active matches \( \lambda \) – defined in (16) – is constant over time. Given that the assignment problem is stationary, the aggregate dynamics of the model closely mirror those of Papanikolaou (2011) and Kogan et al. (2015). In particular, the log aggregate output (12) of the economy equals

\[
\log Y_t = (1 - \phi)x_t + \phi \log K_t,
\]

where \( K_t \) is the quality-adjusted capital stock defined in equation (9). Recall that aggregate labor supply is constant. Equation (23) implies that, at the aggregate level, the model is equivalent to a model with a representative firm, that uses a Cobb-Douglas technology, employs the stock of
quality-adjusted capital $K$, and is subject to a labor-augmenting shock $x$. The law of motion for $K$ is given by

$$dK_t = \left( \lambda e^{\omega_t} \left( \frac{i(\omega_t)}{\lambda} \right)^\alpha - \delta \right) K_t dt,$$

(24)

where the rate of capital accumulation partly depends on the fraction of output devoted to investment

$$i(\omega_t) \equiv \frac{I_t}{Y_t},$$

(25)

which itself is a function of the stationary variable

$$\omega_t \equiv \xi_t + \alpha (1 - \phi) x_t - (1 - \alpha \phi) \log K_t.$$

(26)

The state vector $Z_t = (Y_t, \omega_t)$ is a Markov process that fully characterizes the path of aggregate quantities and prices. The variable $\omega_t$ represents deviations of the current capital stock from its target level – and thus deviations of $Y_t$ from its stochastic trend. As we will see below, the variable $\omega_t$ plays an important role in determining fluctuations in executive pay inequality over medium-run horizons.

The marginal return to new investment $q_t$ plays a key role in our model, which is similar to the role of ‘marginal Q’ (the marginal value of an additional investment) in the neoclassical model. To see this, note that the first-order condition for investment in (11), combined with market clearing, imply that in equilibrium,

$$I_t = \frac{\lambda \alpha}{1 - \alpha} q_t.$$

(27)

The existence of embodied technology shocks implies that the marginal return to new investments $q_t$ is imperfectly correlated with aggregate output $Y_t$ in the short run. Following a positive shock to $\xi$, the return to new investment $q$ rises while output stays constant. Over time, as the economy accumulates more capital $K$, output increases while marginal $Q$ falls. Thus, $q_t$ and $Y_t$ are cointegrated; their ratio is stationary and is proportional to $i(\omega_t)$. This ratio plays an important role for the dynamics of pay inequality.

Next, we examine the model’s predictions about the pay of workers and executives. The equilibrium pay of a production worker is equal to $h w_t dt$. The wage rate of workers engaged in production is determined in equilibrium, and satisfies

$$w_t = (1 - \phi) Y_t.$$

(28)

By contrast, the total level of compensation to the executive currently matched to firm $f$ is given by

$$X_{f,t} = e w_t + \eta (p_{ft} - p^*) \lambda_D q_t.$$

(29)
Examining (29), we see that the flow payment to the executive currently matched to firm $f$ has two components. The first component $e \, w_t$ rewards the executive for her participation in the production process. The second component – which corresponds to $w_{f,t}$ in equation (20) – rewards the executive for her ability to identify new investment opportunities for her firm. This second component of pay is proportional to her value added, which depends on her perceived quality $p_{f,t}$ relative to the quality of an outside hire, the replacement cost, and the net present value of new projects $q_t$.

Equation (29) implies that the average level of executive pay relative to workers will fluctuate over time as a function of the aggregate level of investment opportunities in the economy. We examine two types of inequality: inequality between executives and workers, and inequality across executives.

The disparity in pay between executives and workers, which we define as the ratio of the average level of executive pay to workers, is equal to

$$\frac{\int_0^1 X_{ft} \, df}{h \, w_t} = \frac{e}{h} + \left( \int_0^1 p_{ft} \, df - p^* \right) \frac{\eta \lambda D}{h (1 - \phi)} \frac{q_t}{Y_t}. \quad (30)$$

Examining (30), we see that the fact that executives and workers are endowed with a different amount of effective units of labor services generates a baseline, constant, level of inequality between them. The time variation in the level of inequality between executives and workers is driven by fluctuations in the ratio of the marginal return to new investments $q_t$, and the current aggregate level of output $Y_t$. As we discussed above, this ratio is stationary – it is a monotone function of the state variable $\omega$ that captures the distance between the current level of the capital stock and the technology frontier.

Second, inequality across executives, defined as the cross-sectional standard deviation of log executive pay, can be written as

$$\sigma_t \left( \log \left( e \, w_t + \eta \lambda D \, (p_{ft} - p^*) \, q_t \right) \right) \approx \sigma \left( p_{ft} - p^* \right) \frac{\eta \lambda D}{e (1 - \phi) Y_t + \eta E[p_{ft} - p^*] \lambda D} \, q_t. \quad (31)$$

As before, we see that inequality among executives varies over time as a function of $q_t$ to $Y_t$. To obtain the last approximation in (31), we approximate $X_f$ around its cross-sectional mean, $\int_0^1 X_f \, df$.

In sum, our model generates time variation in inequality – both between executives and workers, and across executives – as a function of $\omega$. Improvements in investment opportunities, as captured by an increase in $q_t$, relative to current output $Y_t$ increase the value of managerial skills for identifying new investment opportunities. As a result, the level of pay of the average executive relative to the earnings of the average worker increases. In our model, the quality of the executive-firm match is the only source of heterogeneity across managers, and it determines the managers’ ability to identify new growth prospects for their firms. Thus, the dispersion in pay across executives also comoves with the level of relative pay.
2 Estimation

Next, we describe how we calibrate the model to the data. We start by providing a very brief overview of the different sources of executive pay data that we use in the paper in Section 2.1. (We discuss this data in more detail in Section 3.) In Section 2.2, we discuss how we choose parameters through a minimum-distance criterion and describe which features of the data help identify the model’s parameters. In Section 2.3, we examine the model’s performance in matching the features of the data that we target, and the resulting parameter estimates. Finally, we provide some insights into the mechanism of our model in Section 2.4.

2.1 Data sources

To exploit time series and cross-sectional variation in executive pay, our empirical analysis is based on a variety of datasets. Our model delivers predictions of our model concern the medium-run dynamics of executive pay inequality in the economy; hence we use the dataset constructed by Frydman and Saks (2010), which provides information on the pay of top executives for most of the twentieth century. Specifically, their data contain the pay of the three highest paid executives in the 50 largest publicly traded corporations in 1940, 1960 and 1990—a total of 101 firms—from 1936 to 2005. This sample is broadly representative of the largest three hundred publicly-traded corporations in each year. For each executive, we use an “ex-ante” measure of total pay, defined as the sum of salary, current bonuses, the payouts from long-term incentive bonuses, and the Black-Scholes value of stock option grants.

A limitation of the Frydman-Saks data is that they cover only a small sample of firms – in a given year, these data contain only about 75 firms on average. Moreover, the sample covers a small number of industries and, within those, the number of firms is usually too small to draw any meaningful conclusions. Thus, these data have limited power to test cross-sectional predictions or to study the variation in executive pay at the industry level. Whenever appropriate, we therefore also evaluate the model using the Execucomp dataset. Execucomp provides information on the pay of top executives in the S&P 500 firms for 1992 and 1993 and, starting in 1994, for all companies included in the S&P 500, S&P MidCap 400, S&P SmallCap 600 indices, as well as some additional firms—covering roughly 1,800 companies each year. For consistency, we restrict the sample to the five highest paid executives in each firm, and we measure total pay for each individual as the sum of salary, current bonus, payouts from long-term incentive bonuses, the value of restricted stock grants, the Black-Scholes value of stock option grants, and other forms of pay—this estimated or ex-ante measure of pay is often called TDC1. Given that both the Frydman-Saks data and Execucomp are based on proxy statements, the definition of executive pay is fairly consistent across samples.

Finally, one of our empirical exercises requires large cross-sectional and industry variation during the 1970s and 1980s—a period in which the investment tax credit (and therefore investment opportunities at the industry level) experienced large, arguably exogenous, variations. For these
tests, we use instead data from Forbes compensation surveys from 1970 to 1991, which contain information on the realized pay for the CEOs of the largest 800 firms in the economy. In contrast to the other two datasets, the Forbes surveys only cover the chief executive and report the gains realized from exercises of stock options instead of their grant values.

To obtain standard measures of firm financial characteristics and firm performance, we match all datasets to Compustat. From the 1930s to the mid-1950s, when Compustat starts, we utilize the financial information hand-collected by Frydman and Saks from the Moody’s Manuals. Only a limited set of variables are available in their data, which limits the controls that we can use in different specifications.

Finally, we utilize various sources of data to construct the statistics that allow us to estimate the parameters in our model. We discuss the construction of these variables in detail in Appendix B.

2.2 Methodology

Our model has a total of 18 parameters. The only parameter that is directly identified by the data is the capital share in production, $\phi$. Following the existing literature, we choose $\phi = 1/3$ (for a textbook reference, see, e.g., Cooley and Prescott, 1995). We estimate the remaining 17 parameters using indirect inference. To do so, we need to choose a set of statistics in the data as estimation targets. The first column of Table 1 reports the 22 statistics that we select as targets. The existing data sources to construct each of these statistics are often only available for a subset of the years covered by our long-run sample on executive pay. To compute these statistics, we use the longest available sample for each of them.

When estimating the model, we make two transformations to the parameters. First, we replace $\lambda_H$ with the difference in match quality between the high and low state, $\lambda_D$. In this manner, we do not need to check that $\lambda_H > \lambda_L$ because we can restrict $\lambda_D$ to be positive. Second, instead of estimating the volatility of the project-specific shock $\sigma_u$, we rewrite the model in terms of its long-run variance, $v_u \equiv \sigma_u^2 / \left( 2 \kappa_u - \sigma_u^2 \right)$. This ensures that the distribution of $u$ has finite variance for all parameter configurations (see Kogan et al., 2015, for more details on the egrodic distribution of $u$).

We next briefly discuss how these statistics help us identify the parameters in our model. Since many of these outcomes are endogenous in our general equilibrium model, the map between parameters and moments is not always straightforward. Thus, we also describe our reasoning for choosing these specific 22 statistics. Visual inspection of the sensitivity of the target moments to the model parameters – the Jacobian matrix $\partial \mathcal{X}(p) / \partial p$, evaluated at the benchmark estimates – confirms our intuition.

The first two moments of aggregate consumption and investment help identify the parameters governing technology shocks, as well as the adjustment cost parameter $\alpha$. The volatility of the risk-free rate, payout growth and the long-run standard deviation of consumption – computed
using the estimator in Dew-Becker (2014) – help pin down the magnitude of the medium-frequency fluctuations in economic growth. Conditional on the parameters determining the equilibrium consumption process, the mean of the risk-free rate also helps identify the subjective discount factor $\rho$.

The elasticity of executive pay to firm size and the average ratio of inequality between executives and workers help identify $\eta$, $c$ and $h$. Cross-sectional differences in investment, Tobin’s $Q$ and executive pay help identify the equilibrium dispersion in firm growth opportunities – a function of dispersion in match quality, which itself depends on $\lambda_D$, $\bar{p}$, $c$ and $\beta$. The elasticity of investment to Tobin’s $Q$ and the serial correlation of investment help identify $\lambda_L$ and $\alpha$. The average tenure of executives helps identify the speed at which the market learns about match quality – which depends on $\lambda_D$ and $\lambda_L$ – as well as the cost of executive termination $c$. The persistence and the cross-sectional dispersion in firm profitability help identify the parameters driving the project-specific shocks $u$.

Our estimation methodology closely follows Kogan et al. (2015). Specifically, we estimate the parameter vector $\theta$ using the simulated minimum distance method (Ingram and Lee, 1991). We denote by $X$ the vector of target statistics in the data and by $\mathcal{X}(\theta)$ the corresponding statistics generated by the model given parameters $\theta$, computed as

$$\mathcal{X}(\theta) = \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(\theta),$$

where $\hat{X}_i(p)$ is the $22 \times 1$ vector of statistics computed in one simulation of the model. We simulate the model at a weekly frequency, and time-aggregate the data to form annual observations. Each simulation has 1,000 firms. For each simulation $i$, we first simulate 100 years of data as ‘burn-in’ to remove the dependence on initial values. We then use the remaining part of that sample, which is chosen to match the longest sample over which the target statistics are computed. Each of these statistics is computed using the same part of the sample as its empirical counterpart. In each iteration we simulate $S = 100$ samples, and simulate pseudo-random variables using the same seed in each iteration.

Our estimate of the parameter vector is given by

$$\hat{\theta} = \arg \min_{\theta} (X - \mathcal{X}(\theta))'W(X - \mathcal{X}(\theta)),$$

where $W = \text{diag}(XX')^{-1}$ is our choice of weighting matrix that ensures that the estimation method penalizes proportional deviations of the model statistics from their empirical counterparts.

9Since dividends are not well defined in our model – the model has no unique dividend policy – we focus instead on net payouts. Depending on the parametrization, net payouts can be negative. Therefore, we target the volatility of the ratio of net payouts to book assets.
We calculate standard errors for the vector of parameter estimates \( \hat{\theta} \) as

\[
V(\hat{\theta}) = \left( 1 + \frac{1}{S} \right) \left( \frac{\partial}{\partial \theta} \mathcal{X}(\theta)'W\frac{\partial}{\partial \theta} \mathcal{X}(\hat{\theta}) \right)^{-1} \frac{\partial}{\partial \theta} \mathcal{X}(\theta)'\mathcal{X}(\theta)W \frac{\partial}{\partial \theta} \mathcal{X}(\theta) \left( \frac{\partial}{\partial \theta} \mathcal{X}(\theta)'W\frac{\partial}{\partial \theta} \mathcal{X}(\theta) \right)^{-1},
\]

(34)

where

\[
\Omega(\hat{\theta}) = \frac{1}{S} \sum_{i=1}^{S} (\hat{X}_i(\hat{\theta}) - \mathcal{X}(\hat{\theta}))(\hat{X}_i(\hat{\theta}) - \mathcal{X}(\hat{\theta}))',
\]

(35)

is the estimate of the sampling variation of the statistics in \( X \) computed across simulations. That is, the standard errors in (34) are computed using the sampling variation of the target statistics across simulations (35). We use (35), rather than the sample covariance matrix. We do so because the statistics that we target are obtained from different datasets (e.g. cross-sectional versus time-series data), and we often do not have access to the underlying data. Since not all of these statistics are moments, computing the covariance matrix of these estimates would be challenging even with access to the underlying data. Under the null of the model, the estimate in (35) would coincide with the empirical estimate. If the model is misspecified, (35) does not need to be a good estimate of the true covariance matrix of \( X \).

Solving each iteration of the model is computationally costly, and thus computing the minimum in (33) using standard numerical methods is computationally infeasible. We therefore use the Radial Basis Function (RBF) algorithm in Björkman and Holmström (2000).

### 2.3 Estimation results

Table 1 presents the quantitative fit of the model to the statistics that we target. Examining columns two to five, we note that our model fits the data reasonably well. The model generates realistic dynamics for aggregate quantities and prices. Further, the model largely replicates the observed dispersion in firm investment, profitability, valuation ratios, and most importantly, executive pay.

Although most of the statistics in simulated data are close to their empirical counterparts, there are a few exceptions. First, and most importantly, the magnitude of the low-frequency fluctuations in pay inequality between executives and workers is substantially smaller in the model than in the data (15% and 50%, respectively). The model is somewhat better at matching fluctuations in inequality across executives (28% in the data vs 15% in the model). Thus, fluctuations in the state of the technology frontier are capable of producing fluctuations in pay inequality that are

---

10 Partly for these reasons, we specify the weighting matrix as \( W = \text{diag}(XX')^{-1} \), rather than scaling by the inverse of the sample covariance matrix of \( X \). In principle we could weight moments by the inverse of (35). However, doing so forces the model to match moments that are precisely estimated but economically less interesting, such as the dispersion in firm profitability or Tobin’s \( Q \).

11 The Björkman and Holmström (2000) algorithm first fits a response surface to data by evaluating the objective function at several hundred points. Then, it searches for a minimum by balancing between local and global search in an iterative fashion. We use a commercial implementation of the RBF algorithm that is available through the TOMLAB optimization package.
approximately one-third to one-half of their realized values. These magnitudes are actually quite substantial if one takes into account that the parameters are fixed over the entire sample period. In other words, the model does not include any other features that may have been contributed to the observed changes in inequality – for instance, changes in the level of taxes, the specificity of executive skills, regulation, the strength of labor unions, the supply of executive talent, and corporate governance. Incorporating some of these forces into the model would undoubtedly lead to more fluctuations in inequality.\(^\text{12}\)

Second, firm investment rates in the model are somewhat less persistent than in the data – investment rates have a serial correlation of 22% in the data and 11% in the model. This result is due to the tight link in the model between executive compensation and investment decisions. We could extend the model to allow for secondary investment decisions – perhaps related to the maintenance of installed assets – that do not require any executive skill. While this modification would allow us to match the persistence in investment rates, it would offer few new insights. Last, the model has difficulty replicating the cross-sectional dispersion in Tobin’s \(Q\) observed in the data; in the model, the inter-quartile range (IQR) in \(Q\) is approximately 40% smaller than the data. Given the fact that part of this dispersion may be measurement error due to imperfections in the empirical measures of the replacement cost of capital, this under-performance of the model is not a major concern.

Table 2 reports the estimated parameters, along with their standard errors. Examining the estimated parameters that govern the level of executive pay, we see that the estimated share of the match-specific surplus that accrues to the executive is roughly one-half. This number is consistent with the estimates of Taylor (2013) obtained using a very different structural model. The two parameters also help determine the level of CEO inequality, \(h\) and \(e\), imply that executives are endowed with approximately 14 times the effective labor input of workers, which helps the model generate an average executive to worker pay ratio of approximately 40. In terms of the composition of executive pay, our parameter estimates imply that the first component of equation (29) (that is, the contribution of the executive to production) accounts, on average, for approximately 42% of her pay. Thus, our parameter estimates imply that executives’ ability to identify new growth opportunities accounts for more than half of their total pay.

Further, we note that the estimated difference in expected project arrival rates across high and low match quality is quite substantial: \(\lambda_D = 0.68\) relative to the mean arrival rate associated with a low quality match \(\lambda_L = 0.05\). This stark difference, which helps the model to fit the mean cross-sectional dispersion in executive pay, implies that learning occurs relatively fast. As a result,

\(^{12}\)For example, Frydman and Molloy (2012) show that an increase in the strength of labor unions may be responsible for the sharp decline in pay disparities between executives and workers during the 1940s. In the context of our model, this would imply that the bargaining power of executives, which we keep constant in our analysis, declined during this period. Similarly, Piketty, Saez, and Stantcheva (2014) suggest that the reduction in income tax rates may have increased executives’ incentives to bargain for higher pay since the 1980s. We could model this as an increase in the bargaining power of executives, which would lead to higher inequality in pay across executives, and between executives and workers.
the model requires fairly high replacement costs \( c = 0.75 \) to prevent low-quality executives to be fired very quickly. Indeed, the fact that executive terminations are relatively infrequent in the data poses a well-known challenge for learning models of executive pay. This difficulty is highlighted by Taylor (2010), who finds that his model requires substantial non-pecuniary costs to shareholders from terminating executives. Since in our model the replacement cost \( c \) does not require real resources – it is mainly a transfer from firms to households– its presence simply helps the model to fit the observed turnover rate.\(^{13}\) Along these lines, the fact that the unconditional probability of drawing a high quality match is fairly low \( \bar{p} = 0.10 \) implies that terminations are rarer than they otherwise would. Further, the estimated likelihood of exogenous turnover is also fairly low, \( \beta = 0.03 \), which helps the model generate substantial dispersion in match quality as well as sufficiently long tenure lengths.

The rest of the parameter estimates are fairly similar to those obtained by Kogan et al. (2015). The volatility of the embodied shock \( \xi \) is approximately twice the volatility of the disembodied shock, implying that vintage effects are especially important in generating time variation in pay inequality. The estimate for the parameter governing decreasing returns to scale in investment is \( \hat{\alpha} = 0.31 \), implying an investment cost function that is not far from quadratic at the project level. Finally, we should emphasize that not all of the parameters are precisely estimated. Their precision reflects the degree to which the output of the model is sensitive to the individual parameter values. For instance, the rate of capital depreciation, the mean values of the two technology shocks \( \mu_x \) and \( \mu_\xi \) and the rate of time preference parameter \( \rho \) are estimated with large standard errors. By contrast, and as it is typically the case, shock volatilities are fairly precisely estimated.

2.4 Examining the model mechanism

To provide further intuition on the model, we briefly examine the forces that lead to pay inequality across executives, as well as variation in inequality over time, within our framework.

We first examine how learning about match quality determines the equilibrium allocation of executives to firms. In Figure 1 we plot four possible paths for the perceived match quality \( p_{f,t} \) (dashed line) along with the actual quality of the match (solid line) over a period of 25 years in simulated data. Shaded regions in the plot correspond to the tenure of different executives.

In the first row, the firm tries a few executives until year 12, all of which prove to be low-quality matches. Following the exogenous separation of the incumbent in year 13, a new executive is hired who was a high-quality match for the firm. Following a sequence of successful investment decisions, shareholders adjust upwards the probability that this match is of high quality. In the second row, a high quality match is hired in year 1 following the dismissal of the previous executive. Within 5 years, shareholders have observed enough successful investment decisions to conclude that the match

\(^{13}\) We could extend the model to allow for other features that would slow down learning, which would likely lower the estimated parameter \( c \). For instance, shareholders could observe investment decisions with noise. We refrain from doing so to keep the number of parameters manageable.
is of high quality. The match is exogenously dissolved in year 21, and the firm unsuccessfully tries to hire a high quality match in following years. In the third row, a low quality match is hired in year 2 following the dismissal of the previous executive, who what revealed to be a low quality match. Even though the new match is also not very successful, a lucky investment decision in the following year implies that it takes another 5 years before shareholders revise their beliefs sufficiently to dismiss the executive. The replacement turns out to be a high quality match, and shareholders realize this relatively quickly. Last, in row four, following the exogenous termination of the incumbent executive in year 4, the firm goes through a sequence of low-quality matches. Even though the executive hired in year 13 is in fact a high quality match to the firm, she fails to make any successful investment decisions and is therefore dismissed in year 15. Her successor – who is also a high quality match – has almost the same fate, but avoids termination right before the threshold in year 17.

The left panel of Figure 2 plots the steady-state distribution of perceived match $p_{f,t}$. Not surprisingly, the distribution is bimodal, with most of the mass being concentrated at the edges. This is the result of two forces. First, the difference in match quality $\lambda_D$ is fairly high, so firms learn about the quality of the match with executives fairly quickly. Second, the exogenous separation rate $\beta$ is fairly low, so matches that are perceived to be high quality are fairly long-lived. This explains the large mass at the top end of the distribution.

We next provide some intuition for the determinants of pay inequality, which are characterized in our model by equations (30) and (31). As we discussed earlier, the disparities in pay between executives and workers are largely the result of performing different functions; both receive compensation for the provision of labor services, but managers are also remunerated for their ability to identify new investment opportunities. The dispersion in executive pay across firms is mainly determined by the cross-sectional dispersion of $p_{f,t}$, as well as the value of investment opportunities $q_t$ relative to the equilibrium wage $w_t$. Since match quality and the return to new investment are complements – they enter multiplicative in equation (29) – pay dispersion among executives increases when there is more dispersion in match quality, or when investment opportunities are greater.

In sum, the dynamics pay inequality across executives and between executives and workers are determined by the ratio of the value of investment opportunities $q_t$ to the equilibrium wage, $w_t$. In our model, this ratio increases monotonically with the mean-reverting state variable $\omega$, as shown in the right panel of Figure 2. Recall that $\omega$, defined in equation (26), represents the transitory deviations of the current (effective) capital stock from its target level. This gap exists because the technology frontier evolves stochastically, whereas the capital stock increases slowly. Thus, $\omega$ can be interpreted as the level of investment opportunities in the economy in a given year.
3 Model predictions

Our next step is to provide direct evidence consistent with the predictions of the model. To
evaluate the model’s quantitative performance, we also compare the magnitude of these empirical
estimates to those performed on simulated data from the model.

3.1 Firm-level evidence

The main prediction of the model is that the level of executive pay in a given firm is closely related
to the value of investment opportunities, as shown in equation (29). We validate this prediction by
examining the association between the level of executive pay in a given firm and various measures of
growth opportunities, and by establishing a positive relationship between executive pay and future
firm growth.

Data description

As we discussed above, depending on the specific analysis, we use three different datasets
containing information on executive pay: the Frydman-Saks data for 1936-2005 (which we refer
(which we refer to as the ‘large panel’). Panels A, B and C of Table 3 presents summary statistics
for executive pay and other firm characteristics for each of these datasets. Except when indicated
otherwise, all our variables are in millions of year 1982 dollars.

The summary statistics reveal some important differences across the three samples. Recall that
the long sample is representative of the 300 largest firms, whereas Forbes is based on 800 large
firms, and Execucomp reports the pay for roughly 1,800 corporations. These differences in sample
selection are accurately reflected on firm size: the average value of log assets is 9.22, 8.22, and
7.29, respectively. Not surprisingly, the Execucomp data have the most dispersion in firm size; the
interquartile differences in log assets are about 2.2, while these differences are much smaller in the
long panel (about 1.6) and in the Forbes data (about 1.4). Despite the difference in the size of firms
across samples, and despite the fact that Forbes covers exclusively CEOs, the mean level of log pay
is lowest in the long sample (about 1.25), and highest in the large panel (about 1.55). The main
reason for this pattern is that executive pay was relatively low until the 1970s, and only began to
increase to the current high levels in the 1980s (see Frydman and Saks, 2010).

All samples, however, reveal a substantial amount of variation in executive pay across firms.
Executives in the 75th percentile of the distribution in pay receive a bit more than double the level
of pay of the executives in the 25th percentile in both the Frydman-Saks and in the Forbes data. In
the Execucomp data, where there is the most variation in firm size, the interquartile differences in
pay are even higher: the 75th percentile in pay is about 3.5 times the level of the 25th percentile in
this sample.
We relate executive pay to the book value of assets, a measure of firm size that reflects the value of assets in place. Not surprisingly, the pay of the average top executive represents a very small fraction of the value of the firm (between 0.02 to 0.21 percent on average across samples). Yet there is large variation in the level of pay relative to firm size across firms and over time; the standard deviation in this measure is as high or higher than the mean in each sample. This simple statistic suggests that there may be large variations in the level of pay across firms, even after controlling for firm size.

In other dimensions, the samples are relatively similar. For the median firm, net income represents about 4 to 5 percent of the value of assets, and stock returns vary between 0.11 and 0.13. The average Tobin’s $Q$ is lower in the Forbes data than in the Execucomp sample, but that is because the data cover different periods (1970-1991 vs 1992-2010). Finally, in the Execucomp sample we can measure the value of investments on physical assets. The average (median) firm has capital expenditures of about 16 (11) percent the value of its existing physical assets.

**Estimating the value of new projects**

The model developed in Section 1 posits a tight link between executive compensation and the value of new investments in the economy $q_t$. While all new projects are ex-ante identical in our theoretical framework, in practice there is substantial heterogeneity in the value of new investment opportunities. Moreover, for a given aggregate realization of $q_t$, there will be variation on whether firms have acquired profitable new projects in a given year. Thus, to assess the predictions of the model using firm-level data, we relate executive pay to a measure of the value of new investments at the firm level.

Constructing an empirical proxy for the marginal value of new investments is not trivial. For example, Tobin’s $Q$, one of the most common measures of investment opportunities in the literature, reflects average rather than marginal values—in other words, Tobin’s $Q$ also reflects the profitability of existing projects. While we return to this and other commonly used measures of growth opportunities in our robustness checks, in our main analysis we instead use a measure of innovation the proxies the marginal value of a firm’s innovative activity in a given year. Specifically, Kogan et al. (2015) – henceforth KPSS – propose that the marginal value of a firm’s innovation can be proxied by the change in stock returns when a firm patents a new idea. Relative to other measures of innovation – such as patent citations – the stock market reaction to patent grants has

---

14 We could extend the model to allow for cross-sectional heterogeneity in the value of new projects, for example by allowing for a match-specific component to the embodied shock $\xi$. If this heterogeneity was persistent across firms, the level of the surplus created by the firm-executive match – and hence the level of executive pay – would vary with it. Allowing for such heterogeneity would likely improve the quantitative performance of the model in matching the data. We choose not to pursue this extension to keep the model parsimonious.

15 Kogan et al. (2012) estimate the net present value of a patent as the change in the dollar value of the firm around a three-day window after the market learns that the firm’s patent application has been successful. Kogan et al. (2012) allow for movements in stock returns around the announcement window that are unrelated to the value of the patent. They construct a filter of the estimated patent value using specific distributional assumptions, and propose a methodology to empirically estimate those parameters using high-frequency data. A limitation of this innovation measure is that it may underestimate the dollar value of a patent by an amount that is proportional to the likelihood
the unique advantage of allowing us to infer the economic – as opposed to the scientific – value of the underlying innovations. We use their measure to construct an estimate of \( q_t \) at the firm-year level, as well as for the entire economy.

We first measure the total dollar value of innovation produced by a given firm \( f \) in year \( t \) by simply summing the estimated values for all patents \( q_j \) that were granted to the firm during that year,

\[
\hat{q}_{f,t} = \sum_{j \in P_{f,t}} q_j,
\]

(36)

where \( P_{f,t} \) denotes the set of patents issued to firm \( f \) in year \( t \). In the context of our model, (36) can be interpreted as the sum of the net present values of all projects acquired by firm \( f \) in year \( t \). Since larger firms tend to produce more patents, we follow KPSS and construct our firm-level innovation measure as

\[
\xi_{f,t} = \frac{\hat{q}_{f,t}}{B_{f,t}},
\]

(37)

where \( B_{f,t} \) is firm size (book assets). We later show that normalizing (37) by the market value of equity leads to very similar results.

While the KPSS measure arguably captures the marginal (as opposed to the average) value of new investment opportunities, it does have some important limitations. First, we have no information on innovations that are not patented. Moreover, we cannot observe the value of patents issued to private firms. Thus, we also analyze the relationship between executive pay and more standard measures of growth opportunities, such as Tobin’s Q and the rate of investment in physical assets.

**Executive pay and firm innovation**

We examine the relationship between executive pay and the value of new investments using the following specification

\[
\log X_{f,t} = a + b \xi_{f,t} + c Z_{f,t} + \varepsilon_{f,t}
\]

(38)

where \( X \) is measured by average executive pay in firm \( f \) in year \( t \) and \( v_{f,t} \) is defined in (37). Depending on the specification, the vector of controls \( Z \) includes log size (measured by book assets, as a way to control for the value of existing assets), profitability (ROA, defined as net income to assets), the firm’s stock return, time, and industry or firm fixed effects. We cluster the standard errors by firm. We also replicate our analysis in data simulated from the model. To facilitate comparisons between the data and the model, we scale \( \xi_{f,t} \) to unit standard deviation.

Equation (38) maps closely to the model. In our framework, executive pay is given by equation (29). The compensation of executives varies over time based on fluctuations in the level output \( Y_t \) and the marginal value of new investments \( q_t \). In the cross-section, the variation in executive pay that the patent application is unsuccessful. This probability is not small; the likelihood that a patent application is successful is approximately 56% in the 1991-2001 period (see Carley, Hegde, and Marco, 2014).
pay is related to the likelihood that the firm-executive match is of high quality $p_{f,t}$ – which would imply that the firm is more likely to acquire new projects. As long as these beliefs are rational, the likelihood of a good-quality match should be related to the *ex-post* realization of the total value of acquired projects that is captured by (36).

We use two samples to estimate equation (38). First, we use the panel data from Frydman and Saks. The long time-series provides more variation over time, including periods of high and low average pay, as well as periods of compression and increased dispersion in inequality. Moreover, the degree of innovative activity as also changed more significantly over the long run. However, a disadvantage of these data is that they only cover a small number of firms and provide very limited industry variation. Thus, we also examine the relationship between pay and investment opportunities using the larger panel based on Execucomp, which covers a shorter time period (1992-2010). To mitigate the impact of outliers, we winsorize the data at the 1 percent level. In the large panel (Execucomp) we compute breakpoints every year, while in the long sample we follow Frydman and Saks (2010) and compute breakpoints over the entire sample period.

We present the results in Table 4. Examining panels A and B, we see that executive pay is strongly related to the marginal value of new projects across specifications and in both samples, even after controlling for firm size. The point estimates are economically significant. Absent any controls, column (1) shows that a one standard deviation increase in $\xi_{f,t}$ is associated with an increase in mean executive pay at the firm level that ranges between 14 percent to 24 percent. After including all controls and fixed effects, the estimated effects are reduced to 4-8 percent. But the magnitude of these estimates is still sizable relative to the cross-sectional dispersion in top executive pay, which ranges from 39 percent to 98 percent, on any given the year in the long sample. Thus, our results indicate that the variation in firm growth opportunities can account for a significant fraction of the heterogeneity in the level of executive pay across firms.

Panel C of Table 4 replicates the analysis in simulated data from the model. We see that the results in simulated data are quantitatively comparable to the data. Since these correlations do not form part of the moments we used to estimate this model, we view them as out of sample evidence supportive of our calibration.

**Executive pay and other measures of growth opportunities**

To further validate our findings, we explore the robustness of our results to two alternative measures of growth opportunities that are more common in the literature: the firm’s realized investment rate (defined as capital expenditures as a ratio of lagged property, plant and equipment—PPE) and the firm’s Tobin’s $Q$ (defined as the sum of the market value of equity plus book value of debt minus inventories and deferred taxes, scaled by lagged PPE). In addition, we also construct an index of firm growth opportunities by taking the first principal component of the KPSS firm innovation measure (scaled by book assets), Tobin’s $Q$ and firm investment rate. The index weights on these three variables are 0.415, 0.682 and 0.603, respectively, and the first principal component
accounts for approximately half of the total variation. Since the Frydman-Saks data do not contain information on investment rates, we perform this analysis only in the Execucomp sample.

Table 5 presents the estimates for equation (38) using these three alternative proxies of $\xi_{f,t}$. Panel B of the Table replicates the analysis in simulated data from the model. Relative to the KPSS measure, the magnitudes of our estimated effects are generally larger, and the estimates are statistically significant. For example, when we include all controls in column (5), we find that a one standard deviation increase in these three measures of growth is associated with an 8 percent to 27 percent increase in executive compensation, which accounts for a significant fraction of the variation in executive pay across firms. Examining Panel B, we see that in two out of the three cases, the estimated coefficients in simulated data are of comparable magnitudes. Focusing on Panel (iii) where we use the first principal component as an index of firm growth opportunities, and on column (5) that includes all controls, a one-standard deviation increase in the growth opportunities index is associated with a 19% increase in compensation with the data, versus a 17% increase in the model. The exception is the coefficient on Tobin’s $Q$, which is sometimes negative or implausibly high. This discrepancy partly reflects the fact that, in both the model and the data, average $Q$ is an imperfect measure of growth opportunities because it also reflects the profitability of existing assets.

Next, we examine the relation between executive pay and an alternative measure of growth opportunities – the investment tax credit (ITC).\footnote{The investment tax credit was first introduced by the Revenue Act of 1962. Until it was eliminated in Tax Reform Act of 1986, the ITC rates changed substantially over time: the tax credit was repealed in 1969, reinstated in 1971, increased in 1975, and modified again in 1981. These changes have often affected investments in different asset classes differentially in rather arbitrary ways. For example, the Tax Reform Act of 1986 changed the investment tax credit available for trucks from 10 percent to 0, while it changed the ITC available for cars from 6 percent to 0.} We follow Cummins, Hassett, Hubbard, Hall, and Caballero (1994) and Edgerton (2010) in using the variation in the tax treatment of different kinds of assets to identify the impact of ITC on investment opportunities. Changes in the effective price the firm pays for new investment should affect the net present value of new investments, but it should not directly impact the profitability of installed capital.\footnote{Indeed, Lyon (1989) finds that changes in firm value are positively related to the expected receipt of investment tax credits, but finds no support for a relation between expected changes in the value of a firm’s existing assets and changes in the firm’s market value.} One advantage of using the ITC is that it provides a plausibly exogenous source of variation for firm growth opportunities. Specifically, there is no clear reason why executive pay should respond to changes in these rates at the industry level other than through their effect on the cost of new investment.\footnote{One possible concern is that changes in ITC may occur at a time in which labor income tax rates (which do directly affect compensation) are also modified. Although this has not always been the case, some tax acts have affected both types of rates. Yet changes in labor income tax rates do not vary by industry in the same manner than ITC does, and so we would expect these changes to be controlled by the time-fixed effects in equation (38). Moreover, Frydman and Molloy (2011) find that executive compensation did not respond much to changes in income tax rates during this period.} Thus, we interpret changes in the ITC as conceptually similar to a shock to $\omega$ in the model that affects firms heterogeneously depending on which industry they belong to.

Since ITC varies at the industry level from 1962 to 1986, the application of this instrument requires a large dataset with sufficient industry variation in pay over this period. Neither the
Frydman-Saks sample nor the Execucomp data are good candidates. Instead, we use the Forbes dataset, which contains realized CEO pay from 1970 to 1991. Table 6 shows the distribution of the computed tax credit over time within the Forbes sample. There is considerable variation in tax credit rates across firm-years: the overall standard deviation of $ITC_{I,t}$ is 2.8%. However, the bulk of that variation comes from the time-series. The cross-sectional standard deviation of $ITC_{I,t}$ across firms ranges from 0.7% in 1970 to 1.1% in 1986 – and it is zero afterwards since the investment tax credit was eliminated in that year.

To examine the impact of changes in firm growth opportunities on executive pay, we re-estimate equation (38) using the investment tax credit in place of $\xi_{f,t}$. We report the estimates in Table 7. We find a statistically significant relation between the investment tax credit rate that applies to firm $f$ in year $t$ and the pay of the firm’s CEO in that year. Our point estimates suggest that a one percentage point increase in the investment tax credit would lead to a 4 percent to 7 percent increase in executive pay, depending on the controls. These results confirm our earlier findings, and suggest that the variation in firm growth opportunities can account for a substantial fraction of the dispersion in executive pay across firms.

Executive pay and future firm growth

Next, we examine the model’s prediction that executive pay should have a positive impact on future firm growth. In our model, the level of executive pay varies across firms as a function of the perceived match quality $p_{f,t}$. Since beliefs about match quality are rational, they should also on average predict future firm growth.

Analyzing the relationship between current levels of pay and future firm growth also allows us to evaluate the predictions of the model without explicitly relying on measures of ex-ante growth opportunities. To assess this relationship, we estimate the impulse response of growth in log firm size $y$ on log executive pay $x$ using local projections (Jorda, 2005). Specifically, we estimate

$$
\log Y_{f,t+s} - \log Y_{f,t} = a_t + a_I + \sum_{l=0}^{L} b^s_l \log X_{f,t-l} + \sum_{l=0}^{L} c^s_l \log Y_{f,t-l} + \varepsilon_{f,t} \tag{39}
$$

The coefficient $b^s_0$ captures the impulse response of $y_{f,t+s}$ on $x_{f,t}$. We choose a lag length of $L = 2$. Our specifications control for time and industry fixed effects. We cluster the errors at the firm level to account for the overlapping observations and serial correlation in firm growth. Moreover, we consider two measures of size – book assets and sales growth – that are also available in the Frydman and Saks (2010) data. To make the connection between the model and the data clear, we also estimate equation (39) in simulated data from the model. To facilitate comparison between the data and the model, we standardize the level of executive pay to unit standard deviation.

We plot the estimated impulse responses $b^s_0$ in Figure 3 for horizons $s$ of 1 to 5 years. Panels A.i and A.ii show the estimated responses in the long sample of Frydman and Saks (2010) and the
large panel of Execucomp, respectively. We find that a one-standard deviation increase in executive pay is associated with a 5 to 10 percent increase in firm size over the next 5 years. Panel B plots the estimated responses in simulated data. The estimates in simulated data are qualitatively and quantitatively similar to the data, which further validates our model.

3.2 Innovation and the dynamics of pay inequality

The empirical results in the previous section document that heterogeneity in firm growth opportunities is economically significantly related to the heterogeneity in the level of executive pay. These findings were identified using firm-level deviations from time trends. Given the attention that aggregate trends in executive pay have attracted in the literature, in this section we change our focus to examine the power of our proposed mechanism for explaining the aggregate fluctuations in inequality in executive pay, both across executives as well as between executives workers. In the next two sections, we consider evidence from the aggregate economy using the long-run series from Frydman and Saks (2010), and cross-industry evidence from Execucomp.

Long-run aggregate evidence

In our model, pay inequality is related to fluctuations in the state variable $\omega$. Intuitively, this variable represents the current level of investment opportunities in the economy – in other words, the distance between the current level of the capital stock and a target level that depends on the two technology shocks $x$ and $\xi$. To examine our model’s ability to account for the aggregate dynamics of pay inequality empirically, we need an empirical proxy for the state variable $\omega$. We use four observable variables that can be mapped directly to $\omega$ in our model. We then obtain their first principal component, to alleviate any concerns that measurement error in any single variable may drive our results.

First, we consider the aggregate estimated value of all patents issued to firms, normalized by either aggregate output,

$$\xi_1^t = \frac{\sum_{f \in F} \hat{q}_ft Y_t}{Y_t}, \quad (40)$$

or aggregate market capitalization

$$\xi_2^t = \frac{\sum_{f \in F} \hat{q}_ft V_t}{V_t}. \quad (41)$$

as in Kogan et al. (2012). The marginal value of new investments, scaled by either output or the value of the stock market, are higher when investment opportunities are good – that is, when the level of the technology frontier $x$ and $\xi$ is far from the current level of capital. We also consider the investment to capital ratio, $IK_t$; this ratio is increasing in $\omega$, since the economy will allocate more resources to investment when the capital stock is far from its target level to reduce this gap.\(^{19}\)

\(^{19}\)More specifically, this ratio is equal to the ratio of investment to the replacement cost of capital, $e^{-\xi_t}K_t$, that is, the cost of replacing the currently installed capital given the technology frontier today. When we construct the equivalent of this ratio in the data, we use the current-cost estimates of non-residential fixed assets as the denominator.
Finally, we also consider the cross-sectional dispersion in firm investment rates, $\sigma_{t(i_{f,t})}$. In the model, increases in the value of new projects magnify the differences in match quality across firms, and hence the observed investment rates.

We denote the vector of observable variables related to $\omega$ by $Z_t = [\xi^1_t, \xi^2_t, IK_t, \sigma_t(i_{f,t})]$. In our model, all four of these variables are stationary. Moreover, Figure 4 shows that they are all monotonically related to the state variable $\omega$ in the model. Motivated by these facts, we use the long-run data on these four measures to construct an empirical proxy for $\omega$. After constructing the vector of variables $Z_t$ as described above, we extract the first principal component (which we refer to as PC1) and use it as our empirical proxy for $\omega$. Since information for some of these series is not available prior to the beginning of Compustat in the mid-1950s, we use probabilistic PCA analysis. The estimated principal component loads positively on all variables in $Z$, and accounts for approximately 74% of the realized variance of the variables in $Z$.

The top row of Figure 5 plots our empirical proxy for $\omega$ (PC1) and the long-run trends in inequality in executive pay using the Frydman-Saks data. The dashed line in the left panel presents the earnings gap between executives and workers, defined as the ratio of the average total executive pay relative to average worker earnings. Overall, this measure of inequality exhibits a J-shaped pattern over the twentieth century: after a sharp decline in the 1940s, the earnings gap continued to compress at a slower rate until the 1970s. Starting in the 1980s, the real level of executive pay has grown at a more rapid pace than the earnings of the average worker. By the end of the sample in the early 2000s, the average top executive earned about 135 times more than the average worker in the economy, about 2.25 times the level of inequality in the late 1930s.

The dashed line in the right panel of Figure 5 shows the cross-sectional dispersion in executive pay across firms, defined as the cross-sectional standard deviation of average firm-level pay in each year. The dispersion in pay was relatively high in the beginning of the sample. Dispersion then declined during the war period, remained stable until the 1970s, and increased substantially after the 1980s. In 2000, the peak of between-executive inequality in our sample, the standard deviation of log executive pay was equal to 1.02, more than twice its level in the 1940s.

In both panels, the solid line presents PC1, our empirical proxy for $\omega$. To provide further validation for this proxy, we note that PC1 lines up well with the three major waves of technological innovation in the U.S. First, our series suggests high values of technological innovation in the 1930s, consistent with the evidence from Field (2003), and Alexopoulos and Cohen (2009, 2011). This measure also indicates that innovation was high during 1960s and early 1970s, as argued by (see, 20)The dispersion in investment rates increases with $\omega$, and it is therefore correlated with medium-run fluctuations in economic growth. This pattern is consistent with the findings of Bachmann and Bayer (2014), who documents that dispersion in investment rates is procyclical.

21Probabilistic PCA allows for missing observations, and estimates the principal component using maximum likelihood in a latent variable model. See Roweis (1998) and Tipping and Bishop (1999) for more details.

22The average level of executive pay declined significantly during War World II. Frydman and Molloy (2012) attribute much of this contraction to the growing power of unions and the decline in the returns to firm size.

23Changes in industry composition over time do not appear to drive this J-shaped pattern. For example, restricting the sample to manufacturing firms leads to very similar trends in average pay over time.
Finally, developments in computing and telecommunication produced an important wave of technological progress in the 1990s and 2000s, which also coincides with high values of PC1.

The top two panels in the Figure reveal that both measures of pay inequality share some common fluctuations with PC1 over the sample period. First, the overall correlation between PC1 and the two pay inequality series are relatively high—about 46 percent and 71 percent, respectively. Moreover, pay inequality and PC1 exhibit similar declines during the 1940s, and increase sharply during the 1990s and 2000s. However, inequality remains mostly flat during the 1960s and 1970s, whereas our proxy for $\omega$ shows an increase in investment opportunities during this period.

The decoupling between investment opportunities and pay inequality at the aggregate level for some periods should not be surprising. Indeed, the U.S. economy witnessed some important structural and institutional transformations during the twentieth century that are outside our model, which likely had a large impact on executive pay and the dispersion in earnings at low frequencies. These forces include, for example, changes in taxes (Frydman and Molloy, 2011; Piketty et al., 2014), the generality of executive skills (Murphy and Zabojnik, 2010; Frydman, 2015), regulation (Murphy, 2013), strength of labor unions (Frydman and Molloy, 2012), changes in the supply of executive talent, and changes in corporate governance (Holmstrom and Kaplan, 2003). Also, our empirical proxies for $\omega$ could be affected by a broadening of the stock market or changes in the composition of the set of publicly traded firms. Our model may therefore be too simple to capture some of the changes in inequality that occur over very low frequencies. Next we analyze whether it has more predictive power for medium-run fluctuations.

To analyze medium-run frequencies, we filter the fluctuations in investment opportunities and in inequality in executive pay using a band-pass filter. Motivated by the work of Comin and Gertler (2006), who show the existence of medium-run cycles that are plausibly related to technological innovation, we restrict attention to frequencies of 5 to 50 years. We plot the filtered series in the bottom row of Figure 5. These medium-run components of executive pay account for approximately one-third of the total variation in pay inequality. We see that the correlation between our proxy for investment opportunities and the two filtered series is fairly substantial – the correlation of our proxy for $\omega$ with the series on executive-worker inequality is 77 percent, and with the cross-sectional dispersion in pay is 78 percent; standard errors computed using Newey-West (20 lags) are 0.065 and 0.11, respectively.

In sum, the aggregate evidence presented here suggests that fluctuations in the level of investment opportunities in the economy can account for a significant share of the fluctuations in the medium-run component of pay inequality, which itself accounts for more than one-third of the total fluctuations

---

24 Indeed, these decades saw important developments in chemicals, oil and computing.
25 Specifically, the medium-run component of inequality between executives and workers has a standard deviation of 14.5%, while the medium-run component of inequality between executives has a standard deviation of 9.6%. These accounts for approximately 30% and 35% of the overall volatility of these two series – see Table 1.
in pay inequality over the 1935 to 2005 period. For comparison, recall that the model can also only account for approximately 30% to 50% of the total fluctuations in inequality – see Table 1.

A shortcoming of this analysis is that the identification is based solely on time-series variation. Though our results show that technological change is one important contributor of these trends in pay inequality, clearly other factors that are not incorporated in our model have also affected the long-run trends in executive pay at the aggregate level. To overcome this shortcoming, next we seek to account for the cross-sectional heterogeneity of pay inequality across industries.

Industry-level evidence

To provide further evidence that technological innovation and managers’ ability to identify new projects are an important component of executive pay inequality, next we examine the relationship between managerial compensation and growth opportunities at the industry level. A main benefit of performing cross-industry comparisons is that it allows us to ‘difference out’ aggregate factors (such as changes in the bargaining power of managers due to changes in the power of labor unions) that affected all industries in a similar manner.

To study disparities in pay between executives and workers across industries, we exploit the larger cross-sectional variation available in the Execucomp data. We measure the average pay of workers at the industry level data across 12 broad sectors using the classifications from the Bureau of Labor Statistics (BLS). We measure inequality as the mean level of the average pay of each firm’s top executive teams in each industry, scaled by the average annual worker earnings in that industry (calculated by multiplying weekly earnings by 52). We measure inequality among executives as the cross-sectional standard deviation of log firm-level executive pay of all firms in industry \( I \) in year \( t \).

We plot the time-series of these two measures of pay inequality for these 12 sectors – in dashed lines – in Figures 6 and 7, respectively. Two main patterns stand out. First, there is considerable time-series variation in both the level and the dispersion of pay across industries. The inequality series in each industry peak at different times for different industries, and the differences in the level of inequality across industries persist over time. Second, the two inequality series have considerable correlation (about 67%) at the industry level. This correlation is not driven by invariant industry characteristics or time effects; removing industry- and year-dummies increases this correlation to 73%, with a \( t \)-statistic of 5.9. Hence, just like in the aggregate data, there is considerable comovement between the level of relative compensation and the dispersion in industry pay across executives.

Next, we analyze whether the sectoral differences in these two measures of executive pay inequality are related to the availability of growth opportunities. As in the previous section, we consider the same four proxies for the current state of investment opportunities in an industry – the vector \( Z \) – as well as the first principal component. The only difference is that, in place of aggregate output, we use firm sales as the denominator of \( \xi_1 \). We compute industry-level equivalents of \( \xi_1 \) as the ratio of the sum of all patents at all firms in the industry divided by the sum of all sales of firms in industry \( I \). We compute \( \xi_2 \) and \( IK_t \) accordingly; \( \sigma(r) \) is computed as the cross-sectional
dispersion of firm investment rates in industry \( I \) in year \( t \). The solid lines in Figures 6 and 7 plot the first principal component of these series, as our preferred industry-specific measure of growth opportunities.

We estimate the following specification,

\[
\log P_{I,t} = a_I + a_t + b G_{I,t} + c \log \bar{Y}_{I,t} + \varepsilon_{I,t},
\]

where \( P_{I,t} \) are measures of pay inequality \( G_{I,t} \) are measures of industry-level growth opportunities; \( a_t \) and \( a_I \) are time and industry dummies, respectively; and \( \bar{Y}_{I,t} \) is the average size of firms in industry \( I \) at time \( t \) measured by sales – we use sales since it appears in the denominator of \( \xi_1 \). To facilitate the comparison of economic magnitudes, we standardize \( G \) and \( \log \bar{Y} \) to unit standard deviation. We cluster standard errors at the industry level.

We report our estimates in Table 8. The estimated effects are statistically significant, and suggest that the relationship between pay inequality and investment opportunities at the industry level is economically meaningful. For example, a one-standard deviation increase in the level of investment opportunities is associated with a 9 to 19 percent increase in the level of inequality between executives and workers (see Panel A). These are sizable magnitudes relative to the standard deviation of (log) inequality, which is 56.9 percent in this sample. Similarly, the results in Panel B show that an increase in one-standard deviation in the industry’s growth opportunities is associated with a 3 to 6 percent increase in the cross-sectional dispersion in executive pay, which accounts for a substantial fraction of its unconditional volatility–about 14.5 percent during that period. Comparing the first and second rows of each panel, we find that the estimated relationship between inequality and growth opportunities is comparable in magnitude to the association between pay inequality and average firm size.

Finally, Table 9 replicates the analysis using instead market capitalization as a measure of firm size. This exercise allows us to examine whether our findings are simply driven by the fact that executive pay is correlated with stock market valuations. A comparison of the coefficients in Tables 8 and 9 reveals that using market value as a measure of firm size has only a minor effect on the estimated coefficient on industry growth opportunities. For example, when we measure investment opportunities using the first principal component across all measures in column (5), the estimated coefficient \( b \) is only 20-25% smaller. Additionally, these results are also inconsistent with a ‘naive’ rent extraction story, under which managers do little else except extracting a constant fraction of firm value. If that were the case, we would expect that the average level of executive compensation in industry \( I \) would mainly be a function of the stock market value of firms in that industry; while clearly market valuations and compensation are related, measures related to the value of investment opportunities clearly play an important role.

We conclude that fluctuations in the level of investment opportunities in the economy can account for a significant fraction of fluctuations in pay inequality, both over time as well as across
industries. This finding complements an extensive literature on executive pay that focuses on the relationship between firm size and compensation. Our analysis suggests that using measures of firm size that reflect only the value of existing assets (such as assets or sales) or that confound the value of existing assets with those of growth opportunities (such as market value), provide only a partial view of the determinants of pay. If one of the main roles of top executives is to identify new projects that foster future firm growth, our paper suggests that executive compensation and pay inequality will depend as well on improvements in technology—a factor that has not been received much attention thus far in the literature on executive pay. Technological change is likely to be one of the main challenges facing large corporations in years to come; our study suggests that it will likely have large impacts on executive pay as well.

4 Conclusion

We argue that the level and the dispersion in executive pay fluctuates following technological improvements, especially when these improvements are embodied in new vintages of capital. We develop a general equilibrium model in which executives add value to the firm not only by participating in production decisions, but also by identifying new investment opportunities. The economic value of these two distinct components of top executives’ jobs varies over time as the technology frontier improves. Improvements in technology that are specific to new vintages of capital raise the skill price of discovering new growth prospects—and thus increase the compensation of executives relative to the pay of workers. When managerial skills primarily consist of their ability to find new projects, dispersion in executive pay across managers and relative to workers will also rise with technological innovation. Improvements in technology that affect all vintages of capital instead increase the value of both skills, as well as worker wages, and thus have a much smaller impact on pay inequality.

Our model delivers testable predictions about the relationship between executive pay and growth opportunities that are quantitatively consistent with the data. We document that (1) executives are paid more in firms with more growth opportunities; (2) medium-run fluctuations in the aggregate time series in executive pay inequality are correlated with aggregate measures of growth opportunities; and (3) industries that have higher growth opportunities also experience an increase in the level and in the dispersion in executive pay.

Our study opens up several avenues for future research. While we posit that top executives have two types of skills, we do not provide any direct evidence that managers vary in their ability to manage assets in place and to identify new valuable projects. A fruitful extension of our analysis would be to identify managerial characteristics that are plausibly related to their ability to identify new investment opportunities for a given firm, and to examine the correlation of these characteristics with executive pay. Further, our model of managerial turnover is fairly stylized. Future work could allow for ex-ante managerial heterogeneity, and examine how the selection of new executives varies
with firm characteristics and with the state of the economy. Finally, our model has little to say about the structure of executive pay. Embedding a moral hazard friction into our setting could deliver rich implications about how managerial incentives should vary with firm characteristics.
Appendix

A. Proofs and Derivations

First, we consider some of the equilibrium relations in order to gain intuition for the overall structure of the solution. Define

\[ \zeta_{j,t} = u_{j,t} e^{\xi_{t}(j)} k_{j,t}. \]  

(A.1)

and

\[ Z_t = \int_{J_t} e^{\xi_{t}(j)} u_{j,t} k_{j,t} dj. \]  

(A.2)

The labor hiring decision is static. The firm managing project \( j \) chooses \( L_{jt} \) as the solution to

\[ \pi_{jt} = \sup_{L_{jt}} \left[ \zeta_{jt} \phi_{jt} (ex_{t} L_{jt})^{1-\phi} - w_{t} L_{jt} \right] \]  

(A.3)

The firm’s choice

\[ L_{jt}^* = \zeta_{jt} \left( (1 - \phi) e^{(1-\phi)x_{t}} \right)^{\frac{1}{\phi}}. \]  

(A.4)

After clearing the labor market, \( \int_{J_t} L_{jt} dj = 1 \), the equilibrium wage is given by

\[ w_{t} = (1 - \phi) e^{(1-\phi)x_{t}} Z_{t}^\phi, \]  

(A.5)

and the choice of labor allocated to project \( j \) is

\[ L_{jt}^* = \zeta_{jt} Z_{t}^{-1}. \]  

(A.6)

Aggregate output of all projects equals

\[ Y_t = \int_{J_t} \zeta_{jt} e^{(1-\phi)x_{t}} Z_{t}^{\phi-1} dj = e^{(1-\phi)x_{t}} Z_{t}^{\phi}. \]  

(A.7)

The project’s flow profits are

\[ \pi_{jt} = \sup_{L_{jt}} \left[ \zeta_{jt} \phi_{jt} (ex_{t} L_{jt})^{1-\phi} - w_{t} L_{jt} \right] = \zeta_{jt} \phi Y_t Z_{t}^{-1} \]  

(A.8)

Because firms’ investment decisions do not affect its own future investment opportunities, each investment maximizes the net present value of cash flows from the new project. Thus, the optimal investment in a new project \( j \) at time \( t \) is the solution to

\[ q_{t} = \sup_{k_{j,t}} \mathbb{E}_{t} \left[ \int_{t}^{\infty} \frac{\Lambda_{s}}{\Lambda_{t}} \pi_{j,s} ds \right] - k_{j,t}^{1/\alpha} = \sup_{k_{j,t}} \left[ P_{t} k_{j,t} e^{\xi_{t}} - k_{j,t}^{1/\alpha} \right], \]  

(A.9)

where \( P_{t} \) is the time-\( t \) price of the asset with the cash flow stream \( \exp(-\delta(s-t))p_{s} \):

\[ P_{t} = \mathbb{E}_{t} \left[ \int_{t}^{\infty} \frac{\Lambda_{s}}{\Lambda_{t}} e^{-\delta(s-t)} p_{s} ds \right]. \]  

(A.10)

The optimal scale of each new project is then given by

\[ k_{t}^* = \left( \alpha e^{\xi_{t}} P_{t} \right)^{\frac{\alpha}{1-\alpha}}. \]  

(A.11)
Note that the solution does not depend on the identity of the firm, i.e., all firms, faced with an investment decision at time \( t \), choose the same scale for the new projects. The optimal investment scale depends on the current market conditions, specifically, on the current level of the embodied productivity process \( \xi_t \), and the current price level \( P_t \).

Using (A.11), we find that the equilibrium value of a new project is equal to

\[
q_t = \alpha^{1/\alpha}(1 - \alpha) \left( e^{\xi_t} P_t \right)^{1/\alpha}. \tag{A.12}
\]

We thus find that the aggregate stock of quality-adjusted installed capital in the intermediate good sector, defined by (9), evolves according to

\[
dK_t = \left( -\delta K_t + \lambda e^{\xi_t} k_t^* \right) dt = \left( -\delta K_t + \lambda e^{\xi_t} \left( \alpha e^{\xi_t} P_t \right)^{1/\alpha} \right) dt, \tag{A.13}
\]

where \( \lambda \) is the aggregate rate of arrival of new projects. In deriving A.13, we have conjectured that \( \lambda \) is constant, a fact we verify below.

An important aspect of (A.13) is that the growth rate of the capital stock \( K_t \) depends only on its current level, the productivity level \( \xi_t \), and the price process \( P_t \). Furthermore, as we show below, we can clear markets with the price process \( P_t \) expressed as a function of the state vector \( X_t = (x_t, \xi_t, K_t) \). Thus, \( X_t \) follows a Markov process in equilibrium.

We express equilibrium processes for aggregate quantities and prices as functions of \( X_t \). For instance, the fact that investment decisions are independent of \( u \) implies that \( Z_t = K_t \). Aggregate investment \( I_t \) is given by

\[
I_t = i(\omega_t) Y_t = \lambda \left( k_t^* \right)^{1/\alpha} = \lambda \frac{\alpha}{1 - \alpha} q_t. \tag{A.14}
\]

where the third equality follows from (A.12). The aggregate consumption process satisfies

\[
C_t = (1 - i(\omega_t)) Y_t = Y_t - I_t = K_t^\phi e^{(1-\phi) x_t} - \lambda \left( k_t^* \right)^{1/\alpha}. \tag{A.15}
\]

Prices of long-lived financial assets, such as the aggregate stock market, depend on the behavior of the stochastic discount factor. In equilibrium, the SDF is determined jointly with the value function of the households, as shown in Lemma ???. Below we fully characterize the equilibrium dynamics and express \( \Lambda_t \) as a function of \( X_t \).

Define the two variables

\[
\chi_t \equiv \log Y_t = (1 - \phi) x_t + \phi \log K_t, \tag{A.16}
\]

and

\[
\omega_t \equiv \xi_t + \alpha (1 - \phi) x_t - (1 - \alpha \phi) \log K_t. \tag{A.17}
\]

\( \omega_t \) and \( \chi_t \) are linear functions of the state vector \( X_t \). In Lemma 1 below, we characterize the SDF and aggregate equilibrium quantities as functions of \( \omega_t \) and \( \chi_t \).

In the formulation of the lemma, we characterize the value function of a household, as well as prices of financial assets, such as \( P_t \) in (A.10), using differential equations. Verification results, such as (Duffie and Lions, 1992, Sec. 4), show that a classical solution to the corresponding differential equation, subject to the suitable growth and integrability constraints, characterizes the value function. Similarly, the Feynman-Kac Theorem (Karatzas and Shreve, 1991, e.g, Theorem 7.6) provides an analogous result for the prices of various financial assets. Because we solve for equilibrium numerically, we cannot show that the classical solutions to our differential equations exist and satisfy the sufficient regularity conditions. With this caveat in mind, in the following lemma we characterize the equilibrium processes using the requisite differential equations.
Lemma 1 (Equilibrium). Let the functions \( i(\omega), v(\omega), g(\omega), h(\omega) \) solve the following system of four ordinary differential equations,

\[
0 = \phi (1 - i(\omega))^{-1} - (\rho + \lambda^{-1}e^\omega i(\omega)^\alpha) v(\omega) + v''(\omega) \frac{1}{2} (\sigma_x^2 + \alpha^2(1 - \phi)^2\sigma_x^2)
+ v'(\omega) \left( \mu_x + \alpha (1 - \phi) \mu_x - (1 - \alpha \phi) \left( \lambda^{-1}e^\omega i(\omega)^\alpha - \delta \right) \right) \\
0 = \frac{i(\omega)}{1 - i(\omega)} \frac{1 - \alpha}{\lambda \alpha} - \rho g(\omega) + g''(\omega) \frac{1}{2} (\sigma_x^2 + \alpha^2(1 - \phi)^2\sigma_x^2)
+ g'(\omega) \left( \mu_x + \alpha (1 - \phi) \mu_x - (1 - \alpha \phi) \left( \lambda^{-1}e^\omega i(\omega)^\alpha - \delta \right) \right) \\
0 = (1 - \phi) (1 - i(\omega))^{-1} - \rho h(\omega) + h''(\omega) \frac{1}{2} (\sigma_x^2 + \alpha^2(1 - \phi)^2\sigma_x^2)
+ h'(\omega) \left( \mu_x + \alpha (1 - \phi) \mu_x - (1 - \alpha \phi) \left( \lambda^{-1}e^\omega i(\omega)^\alpha - \delta \right) \right)
\] (A.18, A.19, A.20)

and the following algebraic equation

\[
\left( \frac{i(\omega)}{\lambda} \right)^{1-\alpha} = \alpha e^{\omega} v(\omega) (1 - i(\omega)). \\
\] (A.21)

Then we can construct price processes and individual policies that satisfy the definition 1, so that \( K_t \) follows

\[
\frac{dK_t}{K_t} = -\delta dt + \lambda e^{\omega t} \left( \frac{i(\omega)}{\lambda} \right)^\alpha dt. \\
\] (A.22)

Proof. We start with a conjecture, which we confirm below, that the equilibrium price process \( P_t \) satisfies

\[
P_t = K_t^{-1} e^{\lambda x t} v(\omega_t) (1 - i(\omega_t)). \\
\] (A.23)

Under this conjecture, the equilibrium aggregate value of assets in place is

\[
V_t \equiv P_t K_t = e^{\lambda x t} v(\omega_t) (1 - i(\omega_t)), \\
\] (A.24)

the value of growth opportunities for the average firm \( (\lambda_f = \lambda) \) is

\[
G_t \equiv \lambda E_t \int_t^\tau \frac{\Lambda_t}{N_t} q_t ds \lambda e^{\lambda x t} g(\omega_t) (1 - i(\omega_t)), \\
\] (A.25)

and the aggregate value of human capital is

\[
H_t = e^{\lambda x t} h(\omega_t) (1 - i(\omega)). \\
\] (A.26)

We then characterize the equilibrium SDF and the optimal policies of the firms and households, and show that all markets clear and the above conjectures are consistent with the equilibrium processes for cash flows and the SDF.

We denote the time-\( t \) net present value of the new projects (the maximum value in (A.12)) by \( q_t \). The aggregate investment process, according to (A.14), is given by

\[
I_t = \lambda \frac{\alpha}{1 - \alpha} q_t. \\
\] (A.27)

Using (A.27) and market clearing (A.14), \( K_t \) follows

\[
\frac{dK_t}{K_t} = \left( -\delta K_t + \lambda e^{\omega t} \left( \frac{\alpha}{1 - \alpha} P_t \right)^{\frac{\alpha}{\alpha - 1}} \right) dt = -\delta dt + \lambda e^{\omega t} \left( \frac{i(\omega_t)}{\lambda} \right)^\alpha dt,
\]
where we have used (A.23), and (A.21) for the last equality. The equilibrium dynamics of the aggregate quality-adjusted capital stock thus agrees with (A.22).

Optimality of household consumption and portfolio choices implies that the SDF
\[ \Lambda_t = e^{-\rho t} C_t^{-1} = e^{-\rho t} e^{-\chi t} (1 - \bar{i}(\omega_t))^{-1}, \] (A.28)
where without loss of generality we have set \( \Lambda_0 = 1 \). Also, it is helpful to define
\[ \pi_t \equiv e^{\rho t} \Lambda_t = e^{-\chi t} (1 - \bar{i}(\omega_t))^{-1}. \] (A.29)

Next, we need to verify that, in equilibrium, the aggregate arrival of projects is constant. Using lemma 3 below, we can write the value of an active match between an executive with perceived talent \( p_{f,t} \) and the firm by
\[ v_t(p_{f,t}) = p_{ft} m_t, \quad \text{where } m_t \equiv \lambda_D E_t \int_1^\tau \frac{\Lambda_s}{\Lambda_t} q_s \, ds, \] (A.30)
where \( \tau \) is the stochastic time the match is dissolved. A firm will terminate its executive if the value of the current match falls below the value of a new match, excluding training costs,
\[ p_{ft} m_t \leq \bar{p} m_t - c \bar{p} m_t \Rightarrow p_{ft} \leq p^* \equiv (1 - c) \bar{p}. \] (A.31)
Since the firing threshold is constant and the stationary distribution of \( p_{f,t} \) exists (see Lemma 2). Hence, the equilibrium level of average executive-firm CEO matches,
\[ \lambda \equiv \lambda_L + \lambda_D \int_{p^*}^1 p f_t(df), \] (A.32)
is a constant.

We are now in a position to complete the proof by verifying that the conjectured price processes in (A.23–A.26) are consistent with the equilibrium SDF above.

**Lemma 2 (Stationary distribution for \( p \)).** The process \( p \) has probability density \( f(p,t) \) that for \( p \in (p^*,1) \) solves
\[ f_t(t,p) = \beta \left( \Delta(\bar{p}) - f_t(t,p) \right) + \lambda_D p (1 - p) f_p(t,p) + (\lambda_L + p \lambda_D) \left( f\left(t, \frac{p \lambda_H}{\lambda_L + p \lambda_D} \right) - f(t,p) \right) \]
where \( \Delta(\bar{p}) \) is the Dirac delta function with point mass at \( \bar{p} \). At \( p = p^* \) the process is reflected to \( \bar{p} \).

Assuming it exists, the stationary density solves the ODE
\[ 0 = \beta \left( \Delta(\bar{p}) - f(p) \right) + \lambda_D p (1 - p) f'(p) + (\lambda_L + p \lambda_D) \left( f\left(p, \frac{p \lambda_H}{\lambda_L + p \lambda_D} \right) - f(p) \right) \]
for \( p \in (p^*,1) \), and once the process hits \( p = p^* \) it is reflected to \( \bar{p} \) and
\[ \lambda = \lambda_L + \lambda_D \int_{p^*}^1 p f(p) \, dp. \] (A.33)

**Proof of Lemma 2.** Lemma follows from a straightforward application of the Kolmogorov Forward equation, taking into account the evolution of \( p \) under rational expectations – equation (14) in the text – the fact that matches are exogenously dissolved at rate \( \beta \) or when \( p \) reaches the firing threshold \( p^* \).
Lemma 3 (Value of Firm-Executive Match). Define the value of a match between a firm and an executive with perceived talent \( p_{ft} \) as

\[
m(p_{ft}, \omega_t, \chi_t) \equiv \pi_{ft} E_t \int_t^\tau \lambda_D \frac{\Lambda_s}{\Lambda_t} q_s \, ds
\]

(A.34)

where \( \tau \) is the earliest of the exogenous or endogenous termination date, whichever comes first. In equilibrium,

\[
m(p_{ft}, \omega_t, \chi_t) = p_{ft} \lambda_D e^{\chi_t} \tilde{g}(\omega_t) (1 - i(\omega_t)),
\]

(A.35)

where \( \tilde{g}(\omega) \) solves the following ODE,

\[
0 = \frac{i(\omega)}{1 - i(\omega)} \frac{1 - \alpha}{\lambda \alpha} - (\rho + \beta) \tilde{g}(\omega) + \tilde{g}''(\omega) \frac{1}{2} \left( \sigma_\rho^2 + \alpha^2 (1 - \phi)^2 \sigma_\phi^2 \right) + \tilde{g}'(\omega) \left( \mu_x + \alpha (1 - \phi) \mu_x - (1 - \alpha \phi) \left( \lambda^{1 - \alpha} e^\omega i(\omega) - \delta \right) \right).
\]

(A.36)

The termination threshold satisfies

\[
p^* = (1 - c) \tilde{p}
\]

(A.37)

Proof. Denote \( \hat{m}(p, \omega, \chi) \equiv \pi m(p, \omega, \chi) \). We know that, in the region \( p \in (p^*, 1) \), that function satisfies the following PDE

\[
0 = p \frac{i(\omega)}{1 - i(\omega)} \frac{1 - \alpha}{\lambda \alpha} - (\rho + \beta) \hat{m}(p, \omega, \chi) - \hat{m}_p(p, \omega, \chi) p (1 - p) \lambda_D
\]

\[
+ \left( \lambda_L + \lambda_D p \right) \left( \hat{m} \left( \frac{p \lambda_H}{\lambda_L + p \lambda_D}, \omega, \chi \right) - \hat{m}(p, \omega, \chi) \right) + D_{\chi, \omega} \hat{m}(p, \omega, \chi)
\]

(A.38)

where optimal firing decisions imply that at the firing threshold \( p^* \) solves

\[
\hat{m}(p^*, \omega, \chi) = (1 - c) \hat{m}(\tilde{p}, \omega, \chi).
\]

(A.39)

Guess that

\[
\hat{m}(p, \omega, \chi) = p \lambda_D \tilde{g}(\omega).
\]

(A.40)

Given our guess, equation (A.38) simplifies to the ODE in (A.36). Given the definition of \( \hat{m} \), we have that

\[
m(p_{ft}, \omega_t, \chi_t) = p_{ft} \lambda_D e^{\chi_t} \tilde{g}(\omega_t) (1 - i(\omega_t)),
\]

(A.41)

The last thing to check is that given our guess, the termination threshold exists and is given by (A.37). Indeed,

\[
p^* \lambda_D e^{\chi_t} \tilde{g}(\omega_t) (1 - i(\omega_t)) - (1 - c) \tilde{p} \lambda_D e^{\chi_t} \tilde{g}(\omega_t) (1 - i(\omega_t)) = 0.
\]

(A.42)

Lemma 4 (Executive Pay). The equilibrium level of executive pay (in excess of their compensation for working in production) is given by

\[
w_{ft} = \eta (p_{ft} - p^*) \lambda_D e^{\chi_t} i(\omega) \frac{1 - \alpha}{\lambda \alpha}
\]

(A.43)

Proof. Define the surplus created from having a current CEO of quality \( p_{ft} \) is

\[
S_{ft} = m(p_{ft}, \omega_t, \chi_t) - (1 - c) m(\tilde{p}, \omega_t, \chi_t)
\]

\[
= (p_{ft} - p^*) \lambda_D e^{\chi_t} \tilde{g}(\omega_t) (1 - i(\omega_t)).
\]

(A.44)
Executives capture fraction $\eta$ of the surplus; hence the present value of CEO pay is $\eta S_{f,t}$. Suppose that the firm promises the CEO a flow rate $w$. Let the CEO’s continuation value be $W_t$; the wage flow satisfies

$$\pi_t W_t = E_t \int_t^\tau e^{-(\rho+\beta)(s-t)} \pi_s w_s ds$$

(A.45)

where $\tau$ is the endogenous termination date. Given our assumptions, we need that this promised value satisfies $W_{f,t} = \eta S_{f,t}$. The fact that the discounted gains process associated with $W_t$ is a martingale implies that

$$0 = \pi w - (\rho + \beta)\pi W + \eta (p - p^*) \lambda D \left[ \tilde{g}''(\omega) \frac{1}{2} \left( \sigma^2 + \alpha^2(1 - \phi)^2 \sigma_x^2 \right) 
+ \tilde{g}'(\omega) \left( \mu_x + \alpha (1 - \phi) \mu_x - (1 - \alpha \phi) \left( \lambda^{1-\alpha} \epsilon^\omega i(\omega)^\alpha - \delta \right) \right) \right]$$

(A.46)

$$0 = \pi w - \eta (p - p^*) \lambda D \left[ \frac{i(\omega)}{1 - i(\omega)} \frac{1 - \alpha}{\lambda \alpha} \right]$$

(A.47)

where in the last equality we have used the fact that $\pi W = \eta \pi S$ and that $\tilde{g}$ solves (A.36). To obtain (A.43) in the text, we can use (A.14) to replace the investment to output ratio $i(\omega_t)$ in (A.43) with the value of investment opportunities $q_t$.

Appendix B: Data

**Aggregate consumption:** We use the Barro and Ursua (2008) consumption data for the United States, which covers the 1834-2008 period. We compute the estimate of long-run risk using the estimator in Dew-Becker (2014). We thank Ian Dew-Becker for sharing his code.

**Aggregate investment and output:** Investment is non-residential private domestic investment. Output is gross domestic product. Both series are deflated by population and the CPI. Data on the CPI are from the BLS. Population is from the Census Bureau. The aggregate investment rate is constructed as the ratio of non-residential investment to the current cost of private non-residential fixed assets (Table 4.1, row 1). Data range is 1927-2010.

**Dividends and payout:** Moments of net payout to assets are from Larrain and Yogo (2008) who use flow of funds data. Data range is 1929-2004. We compute the growth in dividends per share based on the differences between the return with dividends and without dividends of the CRSP value-weighted portfolio, see Hansen, Heaton, and Li (2005), among others, for more details. Data range for dividends per share is 1927-2010.

**Firm Investment rate, Tobin’s Q and profitability:** Firm investment is defined as the change in log gross PPE. Tobin’s Q equals the market value of equity (CRSP December market cap) plus book value of preferred shares plus long term debt minus inventories and deferred taxes over book assets. Firm profitability equals gross profitability (sales minus costs of goods sold) scaled by book capital (PPE). When computing correlation coefficients, we winsorize the data by year at the 1% level to minimize the effect of outliers. We simulate the model at a weekly frequency, $dt = 1/50$ and time aggregate the data at the annual level. In the model, we construct Tobin’s Q as the ratio of the market value of the firm divided by the replacement cost of capital using end of year values. Replacement cost is defined as the current capital stock adjusted for quality, $\hat{K}_{ft} = e^{-\xi_t} K_{ft}$. The investment rate is computed as $I_{ft}/\hat{K}_{ft-1}$, where $I_{ft}$ is the sum of firm investment expenditures in year $t$ and $\hat{K}_{ft-1}$ is capital at the end of year $t - 1$. Similarly, we compute firm profitability as $p_{Z,t} Z_{f,t}/\hat{K}_{f,t-1}$, where $p_{Z,t} Z_{f,t}$ is the accumulated profits in year $t$ and $\hat{K}_{f,t-1}$ is capital at the end of year $t - 1$. Data range is 1950-2010.
**Risk-free rate moments:** We use the reported estimate from the long sample of Barro and Ursua (2008) for the United States and cover the 1870-2008 sample (see Table 5 in their paper). In the data, the risk-free rate is the return on treasury bills of maturity of three months or less. The reported volatility of the interest rate in Barro and Ursua (2008), which equals 4.8%, is the volatility of the realized rate. Hence it is contaminated with unexpected inflation. We therefore target a risk-free rate volatility of 0.7% based on the standard deviation of the annualized yield of a 5-year Treasury Inflation Protected Security (the shortest maturity available) in the 2003-2010 sample. In the model, $r_f$ is the instantaneous short rate; and $R_M$ is the return on the value-weighted market portfolio. We simulate the model at a weekly frequency, $dt = 1/50$ and time aggregate the data at the annual level.

**Executive Compensation** We use the ex-ante value of executive compensation, which includes the Black-Scholes value of granted options. The historical data come from Frydman and Saks (2010). We compute firm-level compensation by computing the average compensation in a given year of the top-3 executives. The large panel data come from Execucomp, where we use tdc1. We compute firm-level compensation by computing the average compensation in a given year of the top-5 executives.

**Investment Tax Credit** We construct industry-level estimates of the effective investment tax credit based on the type of equipment each industry uses. We use data provided by Dale Jorgenson and Jesse Edgerton on investment tax credits for each asset type. We match these variables to the 1982 Capital Flows table from the Bureau of Economic Analysis, which records the amount of investment made by each industry in each asset category. We construct investment tax credit rates at the industry level by taking a weighted average across the assets purchased by each industry, with the weights equal to the percentage of the industry's spending accounted for by each asset.

**Industry-level measures of pay inequality** We construct industry-level estimates of pay inequality using data from Execucomp and the Bureau of Labor Statistics (BLS). Inequality between executives and workers is constructed using the mean (or median) level of firm-level executive pay in each industry, scaled by the average weekly earnings in that industry (times 52). The sectoral definitions are from the BLS, and include Manufacturing (NAICS 31-33), Construction (NAICS 23), Mining and logging (NAICS 21 and 1133), Durable goods (NAICS 321,327,331-337,339), Nondurable goods (NAICS 311-316,322-326), Wholesale trade (NAICS 42), Retail trade (NAICS 44-45), Transportation and warehousing (NAICS 48-49), Utilities (NAICS 22), Information (NAICS 51), Professional and business services (NAICS 54-56), Education and health services (NAICS 61-62), and Leisure and hospitality (NAICS 71-72).
References


Murphy, K. J. and J. Zabojnik (2010, October). Managerial Capital and the Market for CEOs. Working papers, University of South California.


This table reports the fit of the model to the statistics of the data that we target. Growth rates and rates of return are reported at annual frequencies. See main text for details on the estimation method and Appendix B for details on the data construction. We report the mean statistic, along with the 5% and 95% percentiles across simulations. We also report the squared relative deviation of the mean statistic to their empirical counterparts, $Rdev^2_i = (X_i - X_i(p))^2 / X_i^2$. 

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aggregate quantities and prices</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption growth, mean</td>
<td>0.015</td>
<td>0.016</td>
</tr>
<tr>
<td>Consumption growth, volatility</td>
<td>0.036</td>
<td>0.042</td>
</tr>
<tr>
<td>Consumption growth, long-run volatility</td>
<td>0.041</td>
<td>0.053</td>
</tr>
<tr>
<td>Investment, mean share of output</td>
<td>0.089</td>
<td>0.093</td>
</tr>
<tr>
<td>Investment, volatility</td>
<td>0.130</td>
<td>0.098</td>
</tr>
<tr>
<td>Investment and consumption, correlation</td>
<td>0.472</td>
<td>0.384</td>
</tr>
<tr>
<td>Net payout to assets, coefficient of variation</td>
<td>0.575</td>
<td>0.587</td>
</tr>
<tr>
<td>Risk-free rate, mean</td>
<td>0.020</td>
<td>0.015</td>
</tr>
<tr>
<td>Risk-free rate, std</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td><strong>Executive pay and inequality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEO tenure (years), mean</td>
<td>5.800</td>
<td>4.936</td>
</tr>
<tr>
<td>Elasticity of executive pay to firm size</td>
<td>0.350</td>
<td>0.405</td>
</tr>
<tr>
<td>Dispersion in log CEO pay, TS mean</td>
<td>0.553</td>
<td>0.493</td>
</tr>
<tr>
<td>Dispersion in log CEO, log, TS std</td>
<td>0.278</td>
<td>0.152</td>
</tr>
<tr>
<td>Mean CEO pay to worker, log, TS mean</td>
<td>3.708</td>
<td>3.708</td>
</tr>
<tr>
<td>Mean CEO pay to worker, log, TS std</td>
<td>0.496</td>
<td>0.148</td>
</tr>
<tr>
<td><strong>Cross-sectional (firm) moments</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm investment rate, IQR</td>
<td>0.175</td>
<td>0.222</td>
</tr>
<tr>
<td>Firm investment rate, serial correlation</td>
<td>0.223</td>
<td>0.118</td>
</tr>
<tr>
<td>Firm investment rate, correlation with Tobin’s Q</td>
<td>0.237</td>
<td>0.246</td>
</tr>
<tr>
<td>Firm Tobin’s Q, IQR</td>
<td>1.139</td>
<td>0.871</td>
</tr>
<tr>
<td>Firm Tobin’s Q, serial correlation</td>
<td>0.889</td>
<td>0.951</td>
</tr>
<tr>
<td>Firm profitability, IQR</td>
<td>0.902</td>
<td>1.041</td>
</tr>
<tr>
<td>Firm profitability, serial correlation</td>
<td>0.818</td>
<td>0.795</td>
</tr>
</tbody>
</table>

<p>| Distance criterion (mean of Rdev2)             | 0.0629 |</p>
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Estimate</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Executive Labor Market</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive share of match surplus</td>
<td>$\eta$</td>
<td>0.4965</td>
<td>0.4709</td>
</tr>
<tr>
<td>Worker effective supply of labor</td>
<td>$h$</td>
<td>0.0052</td>
<td>0.0006</td>
</tr>
<tr>
<td>Executive excess effective supply of labor</td>
<td>$e$</td>
<td>0.0882</td>
<td>0.0521</td>
</tr>
<tr>
<td>Proportional termination cost</td>
<td>$c$</td>
<td>0.7500</td>
<td>0.0402</td>
</tr>
<tr>
<td>Unconditional match quality</td>
<td>$\bar{p}$</td>
<td>0.0973</td>
<td>0.0253</td>
</tr>
<tr>
<td>Involuntary turnover rate</td>
<td>$\beta$</td>
<td>0.0285</td>
<td>0.0113</td>
</tr>
<tr>
<td><strong>Technology, Preferences and Production</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjective discount rate</td>
<td>$\rho$</td>
<td>0.0015</td>
<td>0.0016</td>
</tr>
<tr>
<td>Decreasing returns to investment</td>
<td>$\alpha$</td>
<td>0.3096</td>
<td>0.0981</td>
</tr>
<tr>
<td>Depreciation rate</td>
<td>$\delta$</td>
<td>0.0300</td>
<td>0.0642</td>
</tr>
<tr>
<td>Disembodied technology growth, mean</td>
<td>$\mu_x$</td>
<td>0.0214</td>
<td>0.1214</td>
</tr>
<tr>
<td>Disembodied technology growth, volatility</td>
<td>$\sigma_x$</td>
<td>0.0768</td>
<td>0.0142</td>
</tr>
<tr>
<td>Embodied technology growth, mean</td>
<td>$\mu_\xi$</td>
<td>0.0014</td>
<td>0.1210</td>
</tr>
<tr>
<td>Embodied technology growth, volatility</td>
<td>$\sigma_\xi$</td>
<td>0.1584</td>
<td>0.0367</td>
</tr>
<tr>
<td>Project mean arrival rate, low quality match</td>
<td>$\lambda_L$</td>
<td>0.0539</td>
<td>0.0265</td>
</tr>
<tr>
<td>Project mean arrival rate, difference high vs low quality match</td>
<td>$\lambda_D$</td>
<td>0.6835</td>
<td>0.0266</td>
</tr>
<tr>
<td>Project-specific productivity, long-run volatility</td>
<td>$v_u$</td>
<td>1.8471</td>
<td>3.6391</td>
</tr>
<tr>
<td>Project-specific productivity, persistence</td>
<td>$\kappa_u$</td>
<td>0.2457</td>
<td>0.0369</td>
</tr>
</tbody>
</table>

This table reports the estimated parameters of the model. When constructing standard errors, we approximate the gradient $\partial \mathcal{X}(\theta)/\partial \theta$ using a five-point stencil centered at the parameter vector $\hat{p}$. 
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exec. Compensation (log 1982 $m)</td>
<td>0.22</td>
<td>0.70</td>
<td>-0.56</td>
<td>-0.28</td>
<td>0.10</td>
<td>0.60</td>
<td>1.18</td>
</tr>
<tr>
<td>Compensation to Assets (%)</td>
<td>0.02</td>
<td>0.02</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>Book assets (log 1982 USDm)</td>
<td>9.22</td>
<td>1.23</td>
<td>7.70</td>
<td>8.42</td>
<td>9.14</td>
<td>10.00</td>
<td>10.80</td>
</tr>
<tr>
<td>Net Income to Assets</td>
<td>0.06</td>
<td>0.05</td>
<td>0.01</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td>Stock Returns</td>
<td>0.10</td>
<td>0.27</td>
<td>-0.25</td>
<td>-0.06</td>
<td>0.12</td>
<td>0.26</td>
<td>0.42</td>
</tr>
<tr>
<td>Firm Innovation, v/B (%)</td>
<td>13.90</td>
<td>22.96</td>
<td>0.00</td>
<td>0.00</td>
<td>4.70</td>
<td>18.11</td>
<td>37.01</td>
</tr>
</tbody>
</table>

A. Long sample (1936-2005)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exec. Compensation (log 1982 $m)</td>
<td>0.33</td>
<td>0.66</td>
<td>-0.42</td>
<td>-0.09</td>
<td>0.27</td>
<td>0.67</td>
<td>1.14</td>
</tr>
<tr>
<td>Compensation to Assets (%)</td>
<td>0.06</td>
<td>0.11</td>
<td>0.01</td>
<td>0.02</td>
<td>0.04</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>Book assets (log 1982 USDm)</td>
<td>8.22</td>
<td>1.10</td>
<td>6.88</td>
<td>7.49</td>
<td>8.18</td>
<td>8.93</td>
<td>9.63</td>
</tr>
<tr>
<td>Net Income to Assets</td>
<td>0.06</td>
<td>0.06</td>
<td>0.01</td>
<td>0.04</td>
<td>0.06</td>
<td>0.09</td>
<td>0.13</td>
</tr>
<tr>
<td>Stock Returns</td>
<td>0.11</td>
<td>0.34</td>
<td>-0.31</td>
<td>-0.07</td>
<td>0.13</td>
<td>0.32</td>
<td>0.50</td>
</tr>
<tr>
<td>Tobin’s Q (log)</td>
<td>0.08</td>
<td>1.06</td>
<td>-1.09</td>
<td>-0.63</td>
<td>-0.05</td>
<td>0.77</td>
<td>1.49</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exec. Compensation (log 1982 $m)</td>
<td>0.44</td>
<td>0.89</td>
<td>-0.68</td>
<td>-0.21</td>
<td>0.37</td>
<td>1.03</td>
<td>1.63</td>
</tr>
<tr>
<td>Compensation to Assets (%)</td>
<td>0.21</td>
<td>0.30</td>
<td>0.02</td>
<td>0.05</td>
<td>0.11</td>
<td>0.24</td>
<td>0.48</td>
</tr>
<tr>
<td>Book assets (log 1982 USDm)</td>
<td>7.29</td>
<td>1.57</td>
<td>5.35</td>
<td>6.15</td>
<td>7.15</td>
<td>8.32</td>
<td>9.52</td>
</tr>
<tr>
<td>Net Income to Assets</td>
<td>0.04</td>
<td>0.12</td>
<td>-0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.09</td>
<td>0.13</td>
</tr>
<tr>
<td>Stock Returns</td>
<td>0.19</td>
<td>0.62</td>
<td>-0.39</td>
<td>-0.14</td>
<td>0.11</td>
<td>0.39</td>
<td>0.78</td>
</tr>
<tr>
<td>Firm Innovation, v/B (%)</td>
<td>13.18</td>
<td>43.33</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>6.91</td>
<td>35.58</td>
</tr>
<tr>
<td>Tobin’s Q (log)</td>
<td>0.97</td>
<td>1.24</td>
<td>-0.51</td>
<td>-0.01</td>
<td>0.85</td>
<td>1.80</td>
<td>2.67</td>
</tr>
<tr>
<td>Investment Rate</td>
<td>0.16</td>
<td>0.19</td>
<td>0.04</td>
<td>0.07</td>
<td>0.11</td>
<td>0.19</td>
<td>0.33</td>
</tr>
</tbody>
</table>

C. Large Panel (1992-2010)

Executive Compensation is the mean compensation of the top-3 managers (in the Frydman and Saks (2010) sample), CEO (in the Forbes data) and top-5 managers (in Execucomp) across firm-years. The compensation data in the long sample is from Frydman and Saks (2010) and includes the ex-ante value of options. The compensation data in the Forbes sample only includes the ex-post (realized) option gains. Compensation data from compustat is TDC1. The value of book assets (Compustat: at) and compensation is deflated to 1982 dollars using the CPI. The firm-level innovation measure $v/B$ is from Kogan et al. (2012). Tobin’s $Q$ is defined as the market value of equity (based on December market capitalization from CRSP) plus the book value of debt (Compustat: dltt) and preferred equity (Compustat: pstkrv) minus inventories (Compustat: invt) and deferred taxes (Compustat: txdb) divided by property plant and equipment (Compustat: ppegt). Industry definitions are based on the SIC2 level. Investment is CAPX (Compustat: capx) scaled by lagged PPE (Compustat: ppegt). Stock returns are from CRSP. firm-year.
Table 4: Executive Pay and Firm Innovation

<table>
<thead>
<tr>
<th>log $X_{f,t}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\nu}<em>{f,t}/B</em>{f,t}$</td>
<td>0.139</td>
<td>0.098</td>
<td>0.059</td>
<td>0.060</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>log $B_{f,t}$</td>
<td>0.302</td>
<td>0.305</td>
<td>0.306</td>
<td>0.317</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>log(1 + ROA$_{f,t}$)</td>
<td>2.353</td>
<td>2.138</td>
<td>2.195</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.356)</td>
<td>(0.362)</td>
<td>(0.309)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(1 + R$_{f,t}$)</td>
<td>0.162</td>
<td>0.166</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.025)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>4883</td>
<td>4883</td>
<td>4881</td>
<td>4869</td>
<td>4869</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.883</td>
<td>0.917</td>
<td>0.921</td>
<td>0.922</td>
<td>0.940</td>
</tr>
</tbody>
</table>

A: Data

i: Long sample (1936-2005)

| log $B_{f,t}$ | 0.407 | 0.405 | 0.407 | 0.370 |      |
|             | (0.005) | (0.005) | (0.005) | (0.012) |      |
| log(1 + ROA$_{f,t}$) | 0.234 | 0.117 | 0.407 |      |      |
|             | (0.047) | (0.047) | (0.045) |      |      |
| log(1 + R$_{f,t}$) | 0.113 | 0.079 |      |      |      |
|             | (0.009) | (0.009) |      |      |      |
| Observations | 28599 | 28599 | 28500 | 28467 | 28467 |
| $R^2$       | 0.243 | 0.598 | 0.600 | 0.603 | 0.769 |

Year FE | Y | Y | Y | Y | Y
Industry (SIC2) FE | Y | Y | Y | - |
Firm FE | - | - | - | Y |

ii: Large panel (1992-2010)

| log $B_{f,t}$ | 0.407 | 0.405 | 0.407 | 0.370 |      |
|             | (0.005) | (0.005) | (0.005) | (0.012) |      |
| log(1 + ROA$_{f,t}$) | 0.234 | 0.117 | 0.407 |      |      |
|             | (0.047) | (0.047) | (0.045) |      |      |
| log(1 + R$_{f,t}$) | 0.113 | 0.079 |      |      |      |
|             | (0.009) | (0.009) |      |      |      |
| Observations | 28599 | 28599 | 28500 | 28467 | 28467 |
| $R^2$       | 0.243 | 0.598 | 0.600 | 0.603 | 0.769 |

Year FE | Y | Y | Y | Y | Y
Industry (SIC2) FE | Y | Y | Y | - |
Firm FE | - | - | - | Y |

B: Model

| log $B_{f,t}$ | 0.414 | 0.414 | 0.414 | 0.601 |      |
|             | (0.108) | (0.108) | (0.106) | (0.132) |      |
| log(1 + ROA$_{f,t}$) | 0.026 | 0.064 | 0.041 |      |      |
|             | (0.055) | (0.058) | (0.043) |      |      |
| log(1 + R$_{f,t}$) | -0.931 | -0.593 |      |      |      |
|             | (0.376) | (0.361) |      |      |      |
| Observations | 30000 | 30000 | 30000 | 30000 | 30000 |
| $R^2$       | 0.217 | 0.465 | 0.465 | 0.467 | 0.721 |

Year FE | Y | Y | Y | Y | Y
Firm FE | - | - | - | Y | Y

Panel A reports estimates of equation (38) in the text. See the notes to Table 3 for variable definitions. Panel B reports estimates of equation (38) in simulated data from the model, using the benchmark parameter estimates reported in Table 2. Standard errors are included in parentheses. In Panel A (data), standard errors are clustered at the firm level. In Panel B (model), standard errors are computed as the standard deviation of parameter estimates across simulations.
Table 5: Executive pay and other measures of growth opportunities

<table>
<thead>
<tr>
<th>log $X_{f,t}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Data (Execucomp)</td>
<td>i. Tobin’s $Q$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log $Q_{f,t}$</td>
<td>0.143</td>
<td>0.247</td>
<td>0.254</td>
<td>0.248</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.202</td>
<td>0.633</td>
<td>0.635</td>
<td>0.635</td>
<td>0.783</td>
</tr>
<tr>
<td>i. Firm Investment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$i_{f,t}$</td>
<td>0.045</td>
<td>0.128</td>
<td>0.128</td>
<td>0.125</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.190</td>
<td>0.606</td>
<td>0.608</td>
<td>0.612</td>
<td>0.775</td>
</tr>
<tr>
<td>iii. Index of Growth Opportunities (PCA)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_{f,t}$</td>
<td>0.195</td>
<td>0.236</td>
<td>0.237</td>
<td>0.231</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.219</td>
<td>0.636</td>
<td>0.637</td>
<td>0.638</td>
<td>0.783</td>
</tr>
<tr>
<td>Observations</td>
<td>27541</td>
<td>27541</td>
<td>27448</td>
<td>27447</td>
<td>27447</td>
</tr>
</tbody>
</table>

B: Model

| log $Q_{f,t}$ | -0.256 | 2.301 | 2.764 | 2.795 | 3.405 |
| | (0.103) | (0.721) | (0.783) | (0.834) | (0.841) |
| $R^2$ | 0.263 | 0.472 | 0.509 | 0.510 | 0.672 |
| i. Firm Investment | | | | | |
| $i_{f,t}$ | 0.207 | 0.190 | 0.190 | 0.197 | 0.152 |
| | (0.021) | (0.019) | (0.018) | (0.027) | (0.023) |
| $R^2$ | 0.269 | 0.426 | 0.426 | 0.427 | 0.583 |
| iii. Index of Growth Opportunities (PCA) | | | | | |
| $G_{f,t}$ | 0.211 | 0.203 | 0.203 | 0.217 | 0.168 |
| | (0.027) | (0.016) | (0.016) | (0.022) | (0.021) |
| $R^2$ | 0.268 | 0.436 | 0.436 | 0.438 | 0.589 |
| Observations | 30000 | 30000 | 30000 | 30000 | 30000 |

Industry FE (Panel A only) | Y | Y | Y | Y | - |
Year FE | Y | Y | Y | Y | Y |
Firm FE | - | - | - | - | Y |
Size | - | Y | Y | Y | Y |
ROA | - | - | Y | Y | Y |
RET | - | - | - | Y | Y |

Table reports estimates of equation (38) in the text using three alternative measures of growth opportunities: Tobin’s $Q$ (panel i), investment (panel ii) and the first principal component across $Q$, investment rate and $v/B$ (panel iii). See the main text and the notes to Table 3 for variable definitions. Industry (SIC2) fixed effects are included in panel A only. Standard errors are included in parentheses. In Panel A (data), standard errors are clustered at the firm level. In Panel B (model), standard errors are computes as the standard deviation of parameter estimates across simulations.
**Table 6**: Distribution of the Investment Tax Credit (ITC) across years

<table>
<thead>
<tr>
<th>year</th>
<th>Mean</th>
<th>SD</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1971</td>
<td>4.5</td>
<td>0.7</td>
<td>3.8</td>
<td>4.3</td>
<td>4.5</td>
<td>5.1</td>
<td>5.1</td>
</tr>
<tr>
<td>1972</td>
<td>4.4</td>
<td>0.7</td>
<td>3.6</td>
<td>4.2</td>
<td>4.5</td>
<td>5.1</td>
<td>5.1</td>
</tr>
<tr>
<td>1973</td>
<td>4.5</td>
<td>0.7</td>
<td>3.6</td>
<td>4.2</td>
<td>4.5</td>
<td>5.1</td>
<td>5.1</td>
</tr>
<tr>
<td>1974</td>
<td>4.5</td>
<td>0.6</td>
<td>3.9</td>
<td>4.3</td>
<td>4.5</td>
<td>5.1</td>
<td>5.1</td>
</tr>
<tr>
<td>1975</td>
<td>6.4</td>
<td>0.9</td>
<td>5.6</td>
<td>6.2</td>
<td>6.5</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td>1976</td>
<td>6.4</td>
<td>0.9</td>
<td>5.6</td>
<td>6.2</td>
<td>6.5</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td>1977</td>
<td>6.4</td>
<td>0.9</td>
<td>5.6</td>
<td>6.1</td>
<td>6.5</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td>1978</td>
<td>6.4</td>
<td>1.0</td>
<td>5.6</td>
<td>6.1</td>
<td>6.5</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td>1979</td>
<td>6.4</td>
<td>1.0</td>
<td>5.6</td>
<td>6.1</td>
<td>6.5</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td>1980</td>
<td>6.5</td>
<td>0.9</td>
<td>5.6</td>
<td>6.1</td>
<td>6.5</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td>1981</td>
<td>7.0</td>
<td>1.0</td>
<td>5.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>1982</td>
<td>6.9</td>
<td>1.0</td>
<td>5.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>1983</td>
<td>6.9</td>
<td>1.0</td>
<td>5.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>1984</td>
<td>6.9</td>
<td>1.0</td>
<td>5.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>1985</td>
<td>6.9</td>
<td>1.1</td>
<td>5.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>1986</td>
<td>6.8</td>
<td>1.1</td>
<td>5.8</td>
<td>6.6</td>
<td>6.9</td>
<td>7.8</td>
<td>7.8</td>
</tr>
<tr>
<td>1987</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1988</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1989</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1990</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1991</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table reports the distribution of the Investment Tax Credit across firms for the period 1970-1991 restricting the sample to those for which we have compensation data from Forbes. We follow Cummins et al. (1994) and construct industry-level estimates of the effective investment tax credit based on the type of equipment each industry uses. We use data provided by Dale Jorgenson and Jesse Edgerton on investment tax credits for each asset type. We match these variables to the 1982 Capital Flows table from the Bureau of Economic Analysis, which records the amount of investment made by each industry in each asset category. We construct investment tax credit rates at the industry level by taking a weighted average across the assets purchased by each industry, with the weights equal to the percentage of the industry's spending accounted for by each asset.
Table 7: Executive pay and growth opportunities: the ITC as a shock to investment opportunities

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log X_{f,t}$</td>
<td>0.062</td>
<td>0.060</td>
<td>0.061</td>
<td>0.062</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>$ITC_{I,t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\log B_{f,t}$</td>
<td>0.246</td>
<td>0.267</td>
<td>0.269</td>
<td>0.293</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>$\log(1 + ROA_{f,t})$</td>
<td></td>
<td></td>
<td></td>
<td>1.750</td>
<td>1.552</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.272)</td>
<td>(0.271)</td>
</tr>
<tr>
<td>$\log(1 + R_{f,t})$</td>
<td></td>
<td></td>
<td></td>
<td>0.160</td>
<td>0.138</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.027)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.575</td>
<td>0.658</td>
<td>0.672</td>
<td>0.673</td>
<td>0.724</td>
</tr>
<tr>
<td>Observations</td>
<td>7498</td>
<td>7498</td>
<td>7498</td>
<td>7477</td>
<td>7385</td>
</tr>
<tr>
<td>Year FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Industry (SIC2) FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>-</td>
</tr>
<tr>
<td>Firm FE</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Y</td>
</tr>
</tbody>
</table>

Table reports estimates of equation (38) using the investment tax credit as an alternative measure of investment opportunities. Sample period is 1970-1991. See notes to Tables 3 and 6 for variable definitions. Standard errors, clustered at the firm level, are included in parentheses.
Table 8: Dynamics of pay inequality: cross-industry evidence

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ξ₁</td>
<td>ξ₂</td>
<td>log I/K</td>
<td>log σ(i)</td>
<td>PC1</td>
</tr>
<tr>
<td>( \log \left( \bar{X}<em>{I,t}/\bar{w}</em>{I,t} \right) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Inequality between executives and workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(mean executive to mean worker)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( G_{I,t} )</td>
<td>0.130</td>
<td>0.090</td>
<td>0.166</td>
<td>0.113</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.022)</td>
<td>(0.045)</td>
<td>(0.028)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>( \log \bar{Y}_{I,t} )</td>
<td>0.344</td>
<td>0.317</td>
<td>0.180</td>
<td>0.268</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>(0.071)</td>
<td>(0.081)</td>
<td>(0.091)</td>
<td>(0.075)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.905</td>
<td>0.895</td>
<td>0.912</td>
<td>0.903</td>
<td>0.917</td>
</tr>
<tr>
<td>B. Inequality among executives</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(dispersion in executive pay)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( G_{I,t} )</td>
<td>0.035</td>
<td>0.029</td>
<td>0.048</td>
<td>0.055</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.008)</td>
<td>(0.022)</td>
<td>(0.014)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>( \log \bar{Y}_{I,t} )</td>
<td>0.081</td>
<td>0.075</td>
<td>0.035</td>
<td>0.053</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.035)</td>
<td>(0.037)</td>
<td>(0.028)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.659</td>
<td>0.652</td>
<td>0.671</td>
<td>0.692</td>
<td>0.689</td>
</tr>
<tr>
<td>Observations</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>I,T</td>
<td>I,T</td>
<td>I,T</td>
<td>I,T</td>
<td>I,T</td>
</tr>
</tbody>
</table>

Table reports estimates of equation (42), where we regress industry level pay inequality on industry-level measures of investment opportunities. In column (1) we use the Kogan et al. (2012) innovation measure at the industry level – which is equal to the sum of the value of patents issued to all firms in industry I at time t scaled by total sales in industry I at time t. In column (2) we scale the value of patents by the market value of all firms in industry I. In column (3) we use the industry level investment rate, constructed as the sum of capital expenditures of all firms in industry I at time t divided by the sum of the capital stock (PPE) of all firms in the industry at time t. In column (4) we use the cross-sectional dispersion of firm investment rates in industry I at time t. In column (5) we use the first principal component of the first three measures. We construct industry-level estimates of pay inequality using data from Execucomp and the Bureau of Labor Statistics (BLS). Inequality between executives and workers (Panel A) is constructed using the mean level of firm-level executive pay in each industry, scaled by the average weekly earnings in that industry (times 52). Inequality among executives is constructed using as the cross-sectional standard deviation of log firm-level executive pay of all firms in industry I at year t. Firm-level pay is the firm-level average of the compensation of the top-5 executives. See Appendix B for the industry definitions. All regressions include controls for the average firm size (sales) in that industry; time (T) and industry (I) dummies. Standard errors are clustered at the industry level and are included in parentheses.
Table 9: Dynamics of pay inequality: cross-industry evidence (controlling for market capitalization)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\xi_1$</td>
<td>$\xi_2$</td>
<td>$\log I/K$</td>
<td>$\log \sigma(i)$</td>
<td>PC1</td>
</tr>
<tr>
<td>$\log (\bar{X}<em>{I,t}/\bar{w}</em>{I,t})$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_{I,t}$</td>
<td>0.046</td>
<td>0.055</td>
<td>0.159</td>
<td>0.099</td>
<td>0.144</td>
</tr>
<tr>
<td>(0.037)</td>
<td>(0.030)</td>
<td>(0.035)</td>
<td>(0.018)</td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>$\log \bar{V}_{I,t}$</td>
<td>0.368</td>
<td>0.383</td>
<td>0.303</td>
<td>0.350</td>
<td>0.317</td>
</tr>
<tr>
<td>(0.045)</td>
<td>(0.045)</td>
<td>(0.042)</td>
<td>(0.033)</td>
<td>(0.038)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.906</td>
<td>0.907</td>
<td>0.927</td>
<td>0.915</td>
<td>0.919</td>
</tr>
<tr>
<td>$\log \sigma_{I,t}(X)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_{I,t}$</td>
<td>0.012</td>
<td>0.020</td>
<td>0.041</td>
<td>0.050</td>
<td>0.050</td>
</tr>
<tr>
<td>(0.011)</td>
<td>(0.008)</td>
<td>(0.018)</td>
<td>(0.011)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>$\log \bar{V}_{I,t}$</td>
<td>0.106</td>
<td>0.109</td>
<td>0.089</td>
<td>0.092</td>
<td>0.087</td>
</tr>
<tr>
<td>(0.042)</td>
<td>(0.040)</td>
<td>(0.039)</td>
<td>(0.037)</td>
<td>(0.036)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.676</td>
<td>0.678</td>
<td>0.697</td>
<td>0.714</td>
<td>0.700</td>
</tr>
<tr>
<td>Observations</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>228</td>
<td>228</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>I,T</td>
<td>I,T</td>
<td>I,T</td>
<td>I,T</td>
<td>I,T</td>
</tr>
</tbody>
</table>

Table reports estimates of equation (42), where we regress industry level pay inequality on industry-level measures of investment opportunities. In column (1) we use the Kogan et al. (2012) innovation measure at the industry level – which is equal to the sum of the value of patents issued to all firms in industry $I$ at time $t$ scaled by total sales in industry $I$ at time $t$. In column (2) we scale the value of patents by the market value of all firms in industry $I$. In column (3) we use the industry level investment rate, constructed as the sum of capital expenditures of all firms in industry $I$ at time $t$ divided by the sum of the capital stock (PPE) of all firms in the industry at time $t$. In column (4) we use the cross-sectional dispersion of firm investment rates in industry $I$ at time $t$. In column (5) we use the first principal component of the first three measures. We construct industry-level estimates of pay inequality using data from Execucomp and the Bureau of Labor Statistics (BLS). Inequality between executives and workers (Panel A) is constructed using as the cross-sectional standard deviation of log firm-level executive pay of all firms in industry $I$ at year $t$. Firm-level pay is the firm-level average of the compensation of the top-5 executives. See Appendix B for the industry definitions. All regressions include controls for the average firm size (equity market capitalization) in that industry; time (T) and industry (I) dummies. Standard errors are clustered at the industry level and are included in parentheses.
Figure 1: Four possible paths for perceived match quality 

Figure plots four possible paths for the perceived match quality $p_{f,t}$ (dashed line), conditional on the true quality of the match $\lambda_{f,t}$ (solid line). The dotted line plots the termination threshold $p^*$. The shaded and non-shaded regions represent the tenure of different executives.
Figure 2: Forces determining pay inequality

Figure plots the steady-state distribution of $p_{f,t}$ in simulated data from the model. We obtain the steady state distribution by simulating 1,000 firms for 1,000 years, and dropping the first half of the sample.
Figure estimates of equation (39) in the long sample (panel A.i), the large panel (panel A.ii) and in simulated data from the model (panel B). Shaded areas correspond to 90% confidence intervals. In Panel A (data), standard errors are clustered at the firm level. In Panel B (model), standard errors are computed as the standard deviation of parameter estimates across simulations.
Figure 4: Model-implied proxies for $\omega$

A. Model

Figure plots the relation between various proxies for $\omega$ and the true state $\omega$ in simulated data from the model (Panel A) and in the data (Panel B). The first two variables are the aggregate estimated value of all patents issued to firms, normalized by either aggregate output ($\xi_1$) or aggregate market capitalization ($\xi_2$) based on Kogan et al. (2012). The third variable is the investment to capital ratio, $I_t/K_t$. Last, the fourth variables is the realized cross-sectional dispersion in firm investment rates, $\sigma_t(i_{f,t})$. 
Figure 5: PC1 of model-implied proxies for $\omega$ vs Pay inequality

Figure plots the relation between our empirical proxy for $\omega$ (the first principal component of the vector $Z_t = [\xi_1^t, \xi_2^t, I_t/Y_t, Q_t, \sigma_i(q_t), \sigma_i(i_{t,t})]$) versus pay inequality, see notes to table 4 for more details). The left panel plots PC1 versus the log ratio of median executive pay to worker; the right panel plots PC1 versus the realized cross-sectional standard deviation of firm-level executive pay. The top panel plots the raw series, while the bottom panel plots the band-pass filtered series, keeping frequencies of 5 to 50 years.
Figure 6: Inequality between executives and workers vs PC1, by sector

Figure plots the time series of industry-level measures of (log) inequality between executives and workers (dashed line) vs the industry-level measures of investment opportunities (PC1). Both series are standardized to zero mean and unit standard deviation. We construct industry-level estimates of pay inequality using data from Execucomp and the Bureau of Labor Statistics (BLS). Inequality between executives and workers is constructed using the mean level of firm-level executive pay in each industry, scaled by the average weekly earnings in that industry (times 52). Firm-level pay is the firm-level average of the compensation of the top-5 executives. Industry definitions are from the BLS; see Appendix B for the industry definitions.
Figure 7: Inequality within executives vs PC1, by sector

Figure plots the time series of industry-level measures of (log) inequality among executives (dashed line) vs the industry-level measures of investment opportunities (PC1). Both series are standardized to zero mean and unit standard deviation. We construct industry-level estimates of pay inequality using data from Execucomp. Inequality among executives is constructed using as the cross-sectional standard deviation of log firm-level executive pay of all firms in industry $I$ at year $t$. Firm-level pay is the firm-level average of the compensation of the top-5 executives. Industry definitions are from the BLS; see Appendix B for the industry definitions.