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

Carola Frydman[†]   Dimitris Papanikolaou[‡]

Abstract

Pay inequality between executives and workers accounts for a substantial fraction of overall pay inequality in the United States. We develop a general equilibrium model that delivers realistic fluctuations in pay inequality – both between executives and workers, and among executives in different firms – 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 in different firms also increases in response to these embodied technological shocks. Our model implies that, controlling for firm size, compensation is higher in fast growing firms. In addition, it implies that pay inequality increases as investment opportunities in the economy improve. Both predictions are consistent with historical and modern data on executive pay.

The dispersion in pay between top executives and workers, as well as among executives in different firms, has fluctuated considerably over the last century. Over this period, pay inequality in the United States has followed a well-documented J-shaped pattern, which largely mirrors movements in overall income inequality (Piketty and Saez, 2003; Frydman and Saks, 2010). Understanding the underlying factors that drive these patterns in executive pay 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 heterogenous 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 executive pay across firms. 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 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 in different firms also comoves with the level of relative pay.

We estimate the parameters of the model using indirect inference. We focus on moments of aggregate investment and consumption, and the dispersion in firm-level investment rates, valuations and profitability. In addition, we also use features of executive pay to estimate the model. 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 both the mean as well as the time-series variation in the disparity of executive pay across firms, and approximately one-half of the observed fluctuations in the executive-to-worker pay ratio. 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 63% of executive pay in the model.

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 constructed by Frydman and Saks (2010), which we extend to the 1936–2014 period. 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. When possible, we present our analysis using both datasets.

We examine the relation between the level of executive pay and firm growth opportunities in two 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 new innovations of Kogan, Papanikolaou, Seru, and Stoffman (2016).[1] Overall, we find a statistically significant and economically substantial association 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 worker wages. 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 (for example, the investment-to-output ratio in the economy). Thus, we create a mapping between observable quantities and the level of inequality that would be predicted by our structural model. The model-implied time series of pay inequality line up well with the empirical series in the beginning and towards the end of the sample. Nevertheless, 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.[2] Specifically, the model does less well in the 1970s and 1960s, as it implies a higher level of pay inequality than the one we observe in the data. The model does better in explaining fluctuations in inequality at medium-run frequencies (frequencies of 5 to 50 years) or changes in inequality over 5-year periods.

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

---

[1] 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.

[2] 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.

replicating the aggregate and cross-sectional dynamics of executive pay inequality. Our framework contains no 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.[3] 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 relations 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 relation 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

---

[3]For 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.

the firm, but also to the expected *growth* in firm size. An important contribution of our study is therefore to help explain some of the variation in executive compensation that remains across firms even after accounting for the current size of the executives' firms. Moreover, 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 relation 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 several 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 comparable conclusions regarding the estimated parameters for the fraction of the match surplus that is captured by executives (about 40 percent). Similar to Taylor (2010), we also find that large firing costs are needed in order to reproduce the level of turnover in the data.

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). Our empirical analysis focuses on the firm's top corporate executives due to the limited availability of information on the compensation of individual workers. We therefore cast our paper as applying to top managers. However, the insights of our model are more broadly applicable. Specifically, the definition of 'executives' in our framework could encompass those workers who have either formal or real authority over firm investment and hiring decisions, or who are responsible for generating new growth opportunities, such as division managers, engineers, research personnel, or key employees in human resources. 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 most valuable talents.

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, Akcigit, 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. Section 1.4 briefly discusses the model's assumptions. 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.[4] 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 capital or 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 all managers provide the same level of effective labor services e, 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

$$J_t = \mathrm{E}_t \left[\int_t^\infty \log(C_s)\,ds \right], \tag{1}$$

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 Λ_t, so the time-t market value of a time-T cash flow X_T is given by

$$\mathrm{E}_t \left[\frac{\Lambda_T}{\Lambda_t} X_T \right]. \tag{2}$$

---

[4]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.

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 ξ, which depends on the state of the technology frontier at the time the project was created; 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

$$y_{j,t} = \left(u_{j,t}\, e^{\xi_{\tau(j)}}\, k_{j,t} \right)^{\phi} \left(e^{x_t}\, L_{j,t} \right)^{1-\phi}, \tag{3}$$

where $L_{j,t}$ is amount of labor allocated to this project and x_t is a shock to the productivity of labor. 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

$$\pi_{j,t} = \sup_{L_{j,t}} \left[\left(u_{j,t}\, e^{\xi_{\tau(j)}}\, k_{j,t} \right)^{\phi} \left(e^{x_t}\, L_{j,t} \right)^{1-\phi} - w_t\, L_{j,t} \right] \tag{4}$$

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, ξ_t and x_t. The shock ξ 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 ξ as an arithmetic random walk

$$d\xi_t = \mu_\xi \, dt + \sigma_\xi \, dB_{\xi,t}, \tag{5}$$

where B_ξ is a standard Brownian motion independent of other shocks in the model. The level of ξ_s denotes the level of technological frontier at time s. Thus, growth in ξ 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

$$dx_t = \mu_x \, dt + \sigma_x \, dB_{x,t}. \tag{6}$$

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

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

$$du_{j,t} = \kappa_u(1 - u_{j,t}) \, dt + \sigma_u \, u_{j,t} \, dB^u_{j,t}, \tag{7}$$

where B^u_j is a standard Brownian motion process independent of B_ξ. We assume that $dB^u_{j,t} \cdot dB^u_{j',t} = dt$ if projects j and j' belong in the same firm f, and zero otherwise. 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 $dB^u_{j,t}$.

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

$$dk_{j,t} = -\delta \, k_{j,t} \, dt, \tag{8}$$

where δ is the economy-wide depreciation rate. Note that the aggregate (quality-adjusted) stock of installed capital in the economy K,

$$K_t = \int_0^1 \left(\sum_{j \in \mathcal{J}_{f,t}} e^{\xi_{s(j)}} k_{j,t} \right) df \tag{9}$$

also depreciates at rate δ.

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, so the level of installed capital in project j equals

$$k_{j,t} = I_{j,t}^{\alpha}, \tag{10}$$

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

$$\nu_t \equiv \sup_{k_{j,t}} \left\{ \mathrm{E}_t \left[\int_t^{\infty} \frac{\Lambda_s}{\Lambda_t} \pi_{j,s} \, ds \right] - k_{j,t}^{1/\alpha} \right\} \tag{11}$$

the net value of a new project implemented at time t under the optimal investment policy, where $\pi_{j,t}$ are profits net of labor costs—defined in equation (4)—and Λ_t is the equilibrium stochastic discount factor defined in Section 1.1. Since all projects created at time t are ex-ante identical, ν 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 sum of the output of all active projects,

$$Y_t = \int_0^1 \left(\sum_{j \in \mathcal{J}_{f,t}} y_{j,t} \right) df. \tag{12}$$

Aggregate output can be allocated to either investment I_t or consumption C_t,

$$Y_t = I_t + C_t. \tag{13}$$

The new investment goods I_t 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 the quality of the match between an executive and a firm $\lambda_{f,t} \in \{\lambda_L, \lambda_H\}$ determines the likelihood that the firm acquires a new investment opportunity at time t. 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 by observing the executive's investment decisions. Standard results on filtering for point processes (Liptser and Shiryaev, 2001) imply that the evolution of their beliefs regarding the match quality $p_{f,t}$ is given by

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

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

$$m_t(p_{f,t}) \equiv p_{f,t} \, E_t \int_t^\tau \lambda_D \frac{\Lambda_s}{\Lambda_t} \nu_s \, ds \qquad (15)$$

where τ is the (stochastic) time at which the match is dissolved, and ν_t is the net present value of a new project created at time t, which is defined in equation (11). Equation (15) 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 $\lambda_L \, dt$.

Firms with poor-quality matches can choose to terminate their executives and replace them with someone new. Specifically, at any point in time, firms will fire the executives if their 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. As 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 probability $\beta \, dt$ each period, the match between CEO and firm is dissolved and the firm must hire a new executive, a process that is explicitly described below. This assumption ensures that the distribution of match quality across executive-firm matches is stationary.

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 $c \, m_t(\bar{p})$; 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 λ_t the mean quality among the current firm-executive matches,

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

Here, λ_t affects the rate at which the economy acquires new projects – or equivalently the rate of capital accumulation. Our modelling assumptions imply that λ_t will be constant over time, hence in what follows we will drop the time subscript. Nevertheless, the equilibrium level of λ 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,

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

We assume that executives can capture a fraction η of that surplus. Importantly, we make the simplifying assumption that the outside option of the executive is zero.[5] 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,

$$W_{f,t} = E_t \int_t^\tau \frac{\Lambda_s}{\Lambda_t} w_{f,s}\, ds \tag{18}$$

and

$$W_{f,t} = \eta\, S_{f,t}. \tag{19}$$

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

$$X_{f,t} = e\, \mathrm{w}_t + w_{f,t}. \tag{20}$$

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.

---

[5]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.

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 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 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 these features, 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.

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 ξ 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.

Fifth, we have assumed that the matching between executives and firms is random. This assumption is in stark contrast with most existing models of executive pay that emphasize

the importance of assortative matching (Gabaix and Landier, 2008; Terviö, 2008). We could change the model to introduce some form of ex-ante heterogeneity among firms that is complementary to the executive's ability to identify profitable new investments. Doing so would magnify the cross-sectional dispersion of pay in our model, but introduce additional technical challenges. In a dynamic model, the relative ranking of firms changes continuously, which, absent any frictions, would imply that executives would need to be reassigned to firms continuously. For this reason, most assignment models of executive pay are essentially static.[6]

Last, we assume that the executive's outside option is zero. This assumption keeps the measure of firms constant over time and 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 $m_t(\bar{p})$, 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 $\{\Lambda_t, \mathrm{w}_t\}$; household consumption decisions $\{C_{i,t}\}$; and firm investment and hiring decisions $\{I_{j,t}, L_{j,t}, p^*\}$ such that given the sequence of stochastic shocks $\{x_t, \xi_t, u_{j,t}, N_{f,t}\}$, $j \in \bigcup_{f \in [0,1]} \mathcal{J}_{f,t}$, $f \in [0,1]$: 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 $w_{f,t}$ satisfies the promise-keeping constraint (18); vi) the labor market clears, $\int_0^1 \left(\sum_{j \in \mathcal{J}_{f,t}} L_{j,t} \right) df = 1$; vii) the demand for new investment equals supply, $\int_0^1 I_{n,t} \, dn = I_t$; viii) the market for consumption clears, and ix) the aggregate resource constraints (12) and (13) are satisfied.*

We next characterize the equilibrium dynamics. To provide some intuition on the inner workings of the model, Figure 1 plots key objects from the model, evaluated using the estimated parameters in Section 2. These plots are qualitatively similar across several parameter configurations. For ease of exposition, we delegate all proofs and derivations to the Online Appendix.

---

[6]A notable exception is Nickerson (2014) who introduces search costs associated with replacing a match. Extending our model along these lines is likely to be a fruitful exercise, but is outside the scope of the current paper.

We begin by showing that the firm's termination threshold p^* is constant across firms and over 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,

$$p_{ft} \leq p^* \equiv (1 - c)\,\bar{p}, \tag{21}$$

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

$$S_{f,t} = (p_{ft} - p^*)\,\lambda_D\,E_t \int_t^\tau \frac{\Lambda_s}{\Lambda_t}\nu_s\,ds, \tag{22}$$

becomes zero.

Panel A of Figure 1 plots the steady-state distribution of perceived match quality $p_{f,t}$. Not surprisingly, the distribution is bimodal, with most of the mass being concentrated at the edges. This is the result of two aspects of the chosen parameters. First, the difference in match quality λ_D is fairly high, so firms learn about the quality of the match with executives fairly quickly. Second, the exogenous separation rate β is fairly low, so matches that are perceived to be high quality are fairly long-lived, implying a large mass at the top end of the distribution.

The fact that the firing threshold p^* is constant implies that the distribution of $p_{f,t}$ across firms is stationary and therefore the mean quality of active matches λ – 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

$$\log Y_t = (1 - \phi)\,x_t + \phi\,\log K_t, \tag{23}$$

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, our model is similar 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

$$\frac{dK_t}{K_t} = \left(\lambda\,e^{\omega_t}\left(\frac{i(\omega_t)}{\lambda}\right)^\alpha - \delta\right)dt, \tag{24}$$

where the rate of capital accumulation partly depends on the distance between the current level of capital and a combination of the two technology shocks x and ξ,

$$\omega_t \equiv \xi_t + \alpha \left(1 - \phi\right) x_t - \left(1 - \alpha \phi\right) \log K_t, \tag{25}$$

as well as the fraction of output devoted to investment

$$i(\omega_t) \equiv \frac{I_t}{Y_t}, \tag{26}$$

which itself is a function of the stationary variable ω.

The state vector (Y_t, ω_t) is a Markov process that fully characterizes the path of aggregate quantities and prices. The variable ω_t represents deviations of the current capital stock from its target level, or equivalently, transitory deviations between investment (or consumption) and output. In a non-stochastic model, ω_t would be constant; in our stochastic model, ω_t is stationary, as we can see from Panel B of Figure 1. Panel C of Figure 1 plots the investment-to-output ratio (26) in our model, which is only a function of ω. This fact, combined with the stationarity of ω, imply that investment, consumption and output are all cointegrated in our model.

The marginal value of new investment ν_t plays a similar role in our model as '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,

$$I_t = \frac{\lambda \alpha}{1 - \alpha} \nu_t. \tag{27}$$

Here, note that the first-order condition for investment in our model is expressed in terms of levels, not ratios, as in the neoclassical model. Equation (27), combined with the fact that the investment-to-output ratio (26) is stationary, implies that the marginal value of new investments ν_t scales proportionally with output Y_t.

Next, we provide some intuition for the determinants of pay inequality in the model. The total level of compensation to the executive currently matched to firm f is equal to

$$X_{f,t} = e \, \mathrm{w}_t + \eta \left(p_{ft} - p^*\right) \lambda_D \, \nu_t. \tag{28}$$

By contrast, the equilibrium pay of a production worker is equal to $h \, \mathrm{w}_t \, dt$. The equilibrium wage rate of workers engaged in production is

$$\mathrm{w}_t = \left(1 - \phi\right) Y_t. \tag{29}$$

As in the neoclassical model with constant labor supply, the wage is exactly proportional to output.

Disparities in pay between executives and workers are largely the result of the differences in the tasks that they perform; both executives and workers receive compensation for the provision of labor services, but only managers are remunerated for their ability to identify new investment opportunities. As a result, the dynamics of pay inequality strongly depend on the ratio of the marginal value of new investment ν_t to the production wage w_t. Panel D of Figure 1 plots this ratio, which is an increasing function of ω.

Panels E and F of Figure 1 present the dynamic response of the value of investment opportunities ν_t, and the equilibrium worker wage w_t to the two technology shocks in the model. As we see in Panel (v), a positive x-shock – the standard TFP shock used in most models – has a symmetric effect on both ν_t and w_t, and thus a negligible effect on the ratio ν/w. Panel (vi) shows that a positive capital-embodied shock ξ leads to a sharp increase in the value of new investments ν_t on impact; by contrast, the equilibrium wage w_t responds with a lag, implying a sharp increase in the ratio ν/w. Over time, as the economy accumulates more capital, output, and hence the production wage increase, whereas the value of new investment falls.

Having developed some intuition for how the skill prices ν and w depend on the state of the economy, we now turn our attention to two measures of pay inequality. First, we are interested in the disparity of pay between executives and workers; the ratio of the average level of executive pay $\bar{X}_t = \int_0^1 X_{ft}\, df$ to workers, is equal to

$$\frac{\bar{X}_t}{h\mathrm{w}_t} = \frac{e}{h} + \frac{\eta\,\mu_p\,\lambda_D}{h}\,\frac{\nu_t}{w_t}, \qquad \text{where} \quad \mu_p \equiv \int_0^1 p_{ft}\, df - p^*. \tag{30}$$

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—captured by the first term. The time variation in the level of inequality between executives and workers is driven by the second term, that is, by fluctuations in the ratio of the marginal return to new investments ν_t, and the current production wage w_t.

Second, we are interested in inequality *across* executives in different firms, defined as the cross-sectional standard deviation of log executive pay, which can be written approximately as,

$$\sigma_t\left(\log\left(X_{f,t}\right)\right) \approx \frac{\eta\,\sigma_p\,\lambda_D\,\nu_t}{e\,\mathrm{w}_t + \eta\,\mu_p\,\lambda_D\,\nu_t}, \qquad \text{where} \quad \sigma_p \equiv \sigma\left(p_{f,t} - p^*\right). \tag{31}$$

As before, we see that inequality among executives varies over time as a function of ν_t/w_t.

Panels G and H of Figure 1 illustrate the dynamic response of the two measures of pay inequality (30) and (31) to the two technology shocks in the model. Consistent with the intuition developed above, a positive disembodied shock x leads to a negligible increase in either measure of inequality. The response of pay inequality to a positive capital-embodied shock ξ is markedly different. A positive shock to ξ leads to a sharp and fairly persistent increase in both measures of pay inequality.

In sum, our model generates time variation in inequality – both between executives and workers, and across executives in different firms. Improvements in investment opportunities, as captured by an increase in ν_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 executive pay across firms also comoves with the level of relative pay. Focusing on magnitudes, we see that in our baseline parametrization, the model can generate not only sizeable, but also fairly persistent fluctuations in inequality: shocks to inequality have a half life of over two decades. The next section examines in detail how these parameters are estimated.

2   Estimation

Next, we describe how we estimate the parameters of the model. We start by providing an overview of the different sources of data that we use in the paper in Section 2.1. In Section 2.2, we discuss how we choose parameters through a minimum-distance criterion. In Section 2.3, we examine the model's performance in matching the features of the data that we focus on, discuss the resulting parameter estimates, and describe which features of the data help identify the model's parameters.

2.1   Data Description

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 regarding 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. For this paper, we extend their data to 2014 using Execucomp.

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. Thus, these data have limited power to test cross-sectional predictions. Whenever appropriate, we evaluate the model using the Execucomp dataset as well. 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. Given that both the Frydman-Saks data and Execucomp are based on proxy statements, the definition of executive pay is fairly consistent across samples.

Our model posits a tight link between executive compensation and the value of new investments in the economy ν_t. Moreover, for a given aggregate realization of ν_t, there will be variation on whether firms have acquired profitable new projects in a given year. Constructing an empirical estimate of the cross-sectional dispersion, as well as persistence, in the acquired value of new projects – which we will term 'firm innovation' – is helpful in both calibrating the model as well as constructing empirical tests of the model mechanism. We use the Kogan et al. (2016) measure as a proxy for the realized value of innovation in a given year. Specifically, Kogan et al. (2016) – henceforth KPSS – propose that the marginal value of a firm's innovation output can be estimated by the change in stock market value when a firm patents a new idea.[7] Relative to other measures of innovation, such as patent citations, the stock market reaction to patent grants has 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 ν_t at the firm-year level, as well as for the entire economy.

---

[7]KPSS 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. KPSS 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.

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

$$\hat{\nu}_{f,t} = \sum_{j \in P_{f,t}} \nu_j, \tag{32}$$

where $P_{f,t}$ denotes the set of patents issued to firm f in year t. In the context of our model, (32) can be interpreted as the sum of the net present values of all projects acquired by firm f in the interval $s \in [t-1, t]$,

$$\nu_{f,t} = \int_{t-1}^{t} \nu_s \, dN_{f,s}. \tag{33}$$

Heterogeneity in project arrival rates – due to variation in the executive-firm match quality – will lead to cross-sectional dispersion in $\nu_{f,t}$.[8] To avoid scale effects, we normalize (32)-(33) with a measure of firm size – either the value of book assets or the firm's market value. Our results are not sensitive to this choice; in our baseline analysis we scale by book assets following KPSS.

To obtain standard measures of firm financial characteristics and 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.

Panels A and B of Table 1 presents summary statistics for executive pay and other firm characteristics for each of the two compensation datasets. Except when indicated otherwise, all our variables are in millions of year 1982 dollars. The summary statistics reveal some important differences across the two samples. Recall that the long sample is representative of the 300 largest firms, whereas 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 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). Despite the difference in the size of firms across samples, the mean level of pay is smaller in the long sample (about 2.1\$m), than the large panel (about 2.5\$m). The main reason for this pattern is that executive pay

---

[8]In our model, dispersion in $\nu_{f,t}$ is driven purely by the extensive margin – the number of projects the firm acquires in a given year. In the data, there is variation in both the extensive as well as the intensive margin – the value of a given patent. 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 ξ, denoted by $\xi_{f,t}$. If $\xi_{f,t}$ 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 $\xi_{f,t}$. 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.

was relatively low until the 1970s, and only began to increase to the current high levels in the 1980s (see Frydman and Saks, 2010). We use the long sample to construct time-series of pay inequality, both between executives and workers, as well as across executives in different firms.

Both samples 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 the Frydman-Saks sample. 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 next 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. We will use the large panel to estimate the elasticity of executive pay to firm size, which is one of the statistics that we use to estimate the model.

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, ϕ. To estimate ϕ, we compute the mean labor share using Flow of Funds data following the methodology of Sekyu and Rios-Rull (2009). Our estimates imply $\phi = 0.325$. We estimate the remaining 17 parameters using indirect inference. To do so, we need to choose a set of statistics in the data that help identify the model parameters.

The first column of Table 2 reports the 21 statistics that we select. We choose three main types of of statistics. First, we choose various moments of the aggregate economy: the first and second moments of investment and consumption growth, the moments of the investment-to-output ratio, the variation in the net payouts, and the mean risk-free rate. Second, we choose statistics related to executive pay and turnover: the median tenure of executives, the elasticity of executive pay to firm size, and the first and second moments of measures of pay inequality – both across executives in different firms as well as between executives and workers. Third, we choose statistics summarizing the dispersion and persistence in firm

investment, innovation and profitability. The Online Appendix discusses the construction of these variables in detail.

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

We estimate the parameter vector θ using the simulated minimum distance method (Ingram and Lee, 1991). Specifically, we choose parameters to minimize the distance between the model and the data,

$$\hat{\theta} = \arg \min_{\theta} \left(X - \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(\theta) \right)' W \left(X - \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(\theta) \right). \tag{34}$$

where $\hat{X}_i$ denotes an estimate of the statistic X in simulation i and W denotes our choice of weighting matrix, which we discuss in more detail below. To construct $\hat{X}$, 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 these statistics are computed. We construct X in exactly the same way in real and simulated data. 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.

We use the optimal weighting matrix,

$$W = \left(\hat{\Sigma} + \frac{1}{S} \Omega(\hat{\theta}) \right)^{-1}, \tag{35}$$

where $\hat{\Sigma}$ is an estimate of the variance-covariance matrix of the statistics X in the data, and

$$\Omega(\hat{\theta}) = \frac{1}{S} \sum_{i=1}^{S} \left(\hat{X}_i(\hat{\theta}) - \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(\theta) \right) \left(\hat{X}_i(\hat{\theta}) - \frac{1}{S} \sum_{i=1}^{S} \hat{X}_i(\theta) \right)' \tag{36}$$

is the estimate of the sampling variation of the statistics in X in simulated data. We estimate Σ by stacking the influence functions of the statistics X following Erickson and Whited (2012) and Davis, Fisher, and Whited (2014). We refer the reader to the Online Appendix for more

details on the estimation of Σ. Similar to two-step GMM, we estimate Ω using a first stage estimate of the model that uses $W = \hat{\Sigma}^{-1}$ as the weighting matrix.

Given the choice of the optimal weighting matrix W in (35), we calculate standard errors for the vector of parameter estimates $\hat{\theta}$ as

$$V(\hat{\theta}) = \left(\hat{G}'W\hat{G}\right)^{-1}, \tag{37}$$

where $\hat{G}$ is an estimate of the gradient of the moment conditions evaluated at the estimated parameter vector $\hat{\theta}$.

2.3   Estimation results

Here, we discuss the estimation results. We begin our discussion by discussing the fit of the model to the data, then discuss our parameter estimates, and last how these parameters are identified by the data.

Model Fit

Table 2 presents the quantitative fit of the model to the statistics that we focus on. We report the empirical statistics (column 1), the mean statistic across model simulations (column 2), and the absolute value of a t-statistic of their difference (column 3). Last, we also report a measure of model fit – the minimized value in (34) – which similar to Hansen's J test is asymptotically χ^2 distributed. See Bazdresch, Kahn, and Whited (2016) for more details.

Comparing the output of the model to the data, we note that in most cases, the economic magnitude of the differences between them is small. The model generates realistic dynamics for aggregate quantities and prices. Further, the model largely replicates the observed dispersion in firm investment, profitability, and most importantly, executive pay. The notable exception is that the magnitude of the low-frequency fluctuations in pay inequality between executives and workers is substantially smaller in the model (24%) than in the data (52%). The model is substantially better at matching fluctuations in inequality across executives in different firms (28% in the data vs 25% in the model). Last, the differences between the model and the data regarding the dispersion and persistence of firm investment rates is statistically significant, even though the magnitudes are economically small, since these statistics are estimated with considerable precision in the data. These deviations imply that the model is formally rejected by the data (p-value is less than 0.01%).

When evaluating the quantitative fit of the model, it is important to keep in mind that the model generates time-series fluctuations in pay inequality between executives and workers using a set of parameters that are fixed over the entire sample period. That is, we do not allow for any other features that may have also contributed to the observed changes in inequality over time, such as, 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. Fluctuations in the state of the technology frontier alone are capable of producing fluctuations in pay inequality between executives and workers that is approximately one-half of the realized value. Incorporating some of these additional forces into the model would undoubtedly lead to more fluctuations in inequality.[9]

Parameter Estimates

Table 3 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 approximately 40%. This number is consistent with the estimates of Taylor (2013) obtained using a very different structural model. The parameters governing the effective labor supply of executives and workers e and h, imply that executives are endowed with approximately 8 times the effective labor input of workers, which helps the model generate an average executive to worker pay ratio of approximately 40. We can also use our parameter estimates to analyze the composition of executive pay. We find that the contribution of executives to production, described by the first component of equation (28), accounts for approximately 37% of the pay for the average executive. Hence, we find that the ability to identify new growth opportunities represents a bit less than two-thirds of average total executive pay. In addition, our estimates imply that, on average, the ratio of executive compensation to firm value is approximately 0.41% in flow terms. This magnitude appears to be reasonable when compared with available data. For example, the average ratio of the total compensation of the five highest-paid executives to the market value of their firms (the numerator in Tobin's Q) was 0.50% for the 1992-2014 period for all companies in the Execucomp dataset.

---

[9]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 estimated difference in expected project arrival rates across high- and low-quality matches is quite substantial ($\lambda_D = 0.69$) 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, the model requires somewhat high replacement costs $c = 0.76$ 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-know 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.[10] Further, our replacement cost is expressed as a fraction of the value of an executive-firm match. Expressed as a fraction of total firm value, the costs of replacing an executive are on average less than 2.5% of the value of the firm in our simulations. This estimate is consistent with Taylor (2010), who estimates that direct termination costs account for approximately 1.6% of firm value. Along these lines, the fairly low ($\bar{p} = 0.07$) unconditional probability of drawing a high quality match implies that terminations are more frequent than they otherwise would. Further, the estimated likelihood of exogenous turnover is also fairly low, $\beta = 0.01$, which helps the model generate substantial dispersion in match quality, as well as sufficiently long tenure.

The rest of the parameter estimates are fairly similar to those obtained by Kogan et al. (2015). The volatility of the embodied shock ξ 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 $\hat{\alpha} = 0.35$, implying an investment cost function that is not too 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, and the precision through which certain moments are estimated. For instance, the rate of capital depreciation δ, the mean values of the two technology shocks μ_x and μ_ξ and the rate of time preference parameter ρ are estimated with relatively large standard errors. By contrast, and as it is typically the case, shock volatilities are fairly precisely estimated. The next section discusses in more detail how these parameters are identified by the data.

---

[10]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.

Identification

We next briefly discuss how the estimated parameters depend on the estimated statistics X. Since ours is a general equilibrium model, the mapping between parameters and moments is not always straightforward. To help us understand which features of the data identify which parameters, we use the Gentzkow and Shapiro (2014) measure of sensitivity, computed as

$$\Lambda = - \left(G'WG\right)^{-1} G'W, \tag{38}$$

where G is an estimate of the gradient of the moment conditions evaluated at the estimated parameter vector $\hat{\theta}$ and W is the weighting matrix (35). To avoid scale effects, we express the sensitivity measure (38) in terms of elasticities. To conserve space, we only summarize the results here, and relegate a more comprehensive description of the matrices G and Λ to the Online Appendix.

The surplus share of executives η is mainly identified by the volatility of the executive-worker pay ratio, and by the mean dispersion in pay across executives in different firms. A higher value of η implies that a larger fraction of executive pay is attributed to growth opportunities. This implies both a more volatile executive-pay ratio and more dispersion in executive pay across firms. The involuntary turnover rate β is identified mainly by the length of the executive's tenure. Longer executive-firm spells, holding all other moments constant, are interpreted as evidence for lower β. The capital depreciation rate δ is identified primarily by the average investment-to-output ratio; in a direct analogue the neoclassical model, a higher depreciation rate of capital implies a higher mean investment-to-output ratio.

The effective labor supply of executives e is primarily identified by the fluctuations in the executive-to-worker pay ratio and in the ratio of payout to assets. Larger values of e imply a higher correlation of pay between executives and workers. Hence a more volatile executive-worker pay ratio is interpreted as evidence for lower e. Further, a higher value of e implies a higher level of executive pay relative to the firm's book value. Thus, a more volatile payout-to-assets ratio is interpreted as evidence for higher e. The effective labor supply of workers h is identified mainly by the mean executive to worker pay ratio. Higher values of h imply a higher level of compensation for production workers, and hence a lower executive-worker pay ratio.

The cost of termination c is mostly identified by three sets of moments: the median length of executive tenure, the mean dispersion in executive pay, and the disparity in innovation outcomes across firms. First, higher values of c imply higher costs of dissolving low quality matches and therefore longer executive spells in the model. Second, higher values of c imply that poor quality matches are maintained for longer, and therefore results in more dispersion

in pay among employed executives. Third, as poor quality matches are maintained for longer, the dispersion in innovation outcomes across firms also increases.

The unconditional probability of a high-quality match $\bar{p}$ is mainly identified by the volatility of the executive-to-worker pay ratio, by the volatility of the investment-to-output ratio, and the median length of tenure across executives. A higher value of $\bar{p}$ implies a higher average quality of active matches. As a result, the average rate λ that the economy accumulates new projects is higher, which, given equations (24)-(25), implies a faster mean-reversion rate for ω, and hence lower a volatility for the investment-to-output ratio $i(\omega)$. The fact that higher $\bar{p}$ implies higher average quality for active matches implies that growth opportunities contribute more to executive pay, which itself implies a more volatile executive-to-work pay ratio. In addition, higher values of $\bar{p}$ imply that a smaller fraction of new firm-executive matches are of low quality, and thus a lower rate of executive termination.

The rate of time preference ρ is primarily identified by the difference between the risk-free rate and the mean rate of consumption growth. To a first-order approximation, the risk-free rate in our model equals the rate of time preference plus the conditional mean of consumption growth. Hence, a larger difference between the average risk-free rate and the mean consumption growth implies a higher value for ρ.

The parameter α governing the decreasing returns to scale in investment is primarily identified by two sets of moments. First, the higher α is, the more responsive investment is to technology shocks. Hence, a higher volatility of investment growth is interpreted as evidence for higher values of α. Second, higher values of α imply that the equilibrium value of a project is smaller, and thus a smaller fraction of executive pay comes from growth opportunities. As a result, the difference between the volatility of investment-to-output ratio (which is increasing in α) and the volatility of the executive-pay ratio (which is decreasing in α) also helps identify this parameter.

The project arrival rate for low quality matches λ_L is primarily identified by the elasticity of executive pay to firm size, and the volatility of the executive-to-worker pay ratio. The higher is the sensitivity of executive pay to firm size, the tighter is the link between firm growth and firm pay. Holding all the other moments constant, a higher elasticity of firm size to executive pay will imply a lower value for λ_L. Further, higher values of λ_L imply that more new projects are created per period, which in turn leads to less capital being allocated to any given project. This in turn implies a lower equilibrium value per project and thus lower dispersion in executive pay across firms. The difference in project arrival rates across high- and low-quality matches λ_D is mainly identified by the median length in executive tenure, the dispersion in innovation outcomes across firms and the mean dispersion in pay

across executives. First, an increase in λ_D implies faster learning about match quality and hence faster dissolutions of low-quality matches. Second, higher values of λ_D imply larger differences in the rate of innovation across firms. Third, as λ_D increases, pay differences across executives are magnified, implying larger cross-sectional dispersion in pay.

The volatility of disembodied shocks σ_x is primarily identified by the joint distribution of investment and consumption growth and the volatility of the executive-to-worker pay ratio. Specifically, the shock x affects investment and consumption symmetrically. A higher volatility of consumption or investment growth – or a higher correlation between the two – is interpreted by the model as evidence for higher σ_x. Further, x affects executive and worker pay fairly symmetrically. Holding the other moments constant, a more volatile worker-pay ratio is interpreted as evidence for lower σ_x.

The volatility of the embodied shock σ_ξ is mainly identified by three sets of moments. First, the ξ shock affects investment more than consumption on impact. Therefore, a more volatile investment growth, or investment-to-output ratio, is interpreted as evidence for higher σ_ξ. Second, ξ affects executive pay on impact, but worker pay with delay. Thus, a higher volatility in the executive-pay ratio is interpreted by the model as suggestive of higher σ_ξ. Third, ξ affects the optimal scale of investment much more than the value of innovation across firms. Thus, a higher dispersion in investment rates, relative to the dispersion in innovation rates, implies that ξ has to be more volatile.

The mean of the disembodied shock μ_x is primarily identified by the mean of consumption growth, though other moments seem to also have an effect. The mean of the embodied shock μ_ξ is identified by average consumption growth and the moments of the investment-to-output ratio. Higher values of μ_ξ imply that the mean of the stationary distribution of ω is higher, which implies higher investment-to-output. It is important to note that the means of the two technology processes, μ_x and μ_ξ have similar impact on model quantities, so it is hard to identify them separately. This pattern, along with the fact that the mean of consumption growth is relatively imprecisely estimated, results in relatively large standard errors for μ_x and μ_ξ. Last, the parameters governing the process for idiosyncratic productivity v_u and κ_u are identified by the dispersion and persistence of profitability across firms, respectively.

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. These correlations

do not form part of the moments we used to estimate this model, hence can be interpreted as 'out of sample' tests of 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 (28). 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 documenting a positive relation between executive pay and future firm growth.

Executive pay and firm innovation

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

$$\log X_{ft} = a + b\, \frac{\hat{\nu}_{f,t}}{B_{f,t}} + c\, Z_{ft} + \varepsilon_{ft} \tag{39}$$

where $X_{f,t}$ is average executive pay in firm f in year t, $B_{f,t}$ is firm size (book assets) and $\hat{\nu}_{f,t}$ is the KPSS estimate of the dollar value of innovation to firm f in year t, defined in (32). Depending on the specification, the vector of controls $Z_{f,t}$ 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, year dummies, and industry or firm fixed effects. In the Online Appendix, we show that our results are robust to using alternative measures of firm size – the market value of the firm or sales. We cluster the standard errors by firm and year. To facilitate comparisons between the data and the model, we scale $\hat{\nu}_{f,t}/B_{f,t}$ to unit standard deviation.

Equation (39) maps closely to the model. In our framework, executive pay is given by equation (28). The compensation of executives varies over time based on fluctuations in the level output Y_t and the marginal value of new investments ν_t. In the cross-section, the variation in executive 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 (33).

We use two samples to estimate equation (39). First, we use the panel data from Frydman and Saks, extended to 2014. 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 has also changed more significantly over the long run. However, a disadvantage of these data is that they only cover a small number of firms. Thus, we also examine the relation between pay and investment opportunities using the larger panel based on Execucomp, which covers a shorter time period (1992-2014). 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 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 $\hat{\nu}_{f,t}/B_{f,t}$ 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 5-7 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.

The patterns documented thus far are not dependent on the particular measure of firm size. Tables A.1 and A.2 in the Online Appendix reproduce Table 4, but they use instead the market value of the firm (defined by the sum of the market value of the firm's equity and the book value of its debt) or the value of sales, respectively. The estimated effects of our main variable of interest $\hat{\nu}_{f,t}/B_{f,t}$ are robust to using these alternative measures of firm size. In most cases, the estimated coefficients b are typically somewhat larger in magnitude. Our point estimates of b only become smaller when we use market values in the Execucomp data, but they remain statistically significant.

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 and the firm's Tobin's Q. In addition, we also construct an index of firm growth opportunities by taking the first principal component of the KPSS firm innovation measure, Tobin's Q and the firm's investment rate. The index weights on these three variables are 0.42, 0.68 and 0.60, 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 (39) using these three alternative proxies in the place of $\hat{\nu}/B$. 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 opportunities is associated with an 8 percent to 26 percent increase in executive compensation, which accounts for a significant fraction of the variation in executive pay across firms. These patterns are robust to using market value (as shown in Table A.3 in Online Appendix) or sales (as shown in Table A.4 in Online Appendix) as alternative measures of firm size. The estimated effects do become quantitatively smaller when we use market values, but they are statistically significant across specifications.

Executive pay and growth opportunities in simulated data

Next, we replicate the empirical analysis in the previous two sections using data simulated from the model. In Table 6 we report the mean coefficients across 100 model simulations, as well as the standard deviation of the point estimates across simulations. We see that the results in simulated data are in most cases quantitatively comparable to the data. Focusing on Panel (D), in which 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 18% increase in compensation with the data, versus a 17% increase in the model.

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 model. The exception is the coefficient on Tobin's Q, which is sometimes negative or implausibly high in simulated data. 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.

Last, our model can also largely replicate these patterns using our alternative definitions of firm size (see Tables A.5 and A.6 in the Online Appendix). Consistent with the data, the point estimates of b become smaller if we measure firm size by the market value of the firm when estimating equation (39) in simulated data. In the model, this happens because firm market values include both the value of the firm's current assets and also the firm's future growth opportunities, which reduces the explanatory power of the measures of investment opportunities used to estimate (39).

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 relation between current levels of pay and future firm growth allows us to evaluate the predictions of the model without explicitly relying on ex-ante measures of growth opportunities.

To assess this correlation, we estimate the impulse response of growth in log firm size ($\log Y_{f,t}$) on log executive pay ($\log X_{f,t}$) using local projections (Jorda, 2005). Specifically, we estimate

$$\log Y_{f,t+s} - \log Y_{f,t} = a_t + a_I + \sum_{l=0}^{L} b_l^s \log X_{f,t-l} + \sum_{l=0}^{L} c_l^s \log Y_{f,t-l} + \varepsilon_{f,t} \qquad (40)$$

The coefficient b_0^s captures the impulse response of log size on log pay. We choose a lag length of $L = 2$. Our specifications control for time and industry fixed effects. We cluster the errors at the firm and year level to account for the overlapping observations and serial correlation in firm growth, as well as correlation in firm growth rates in a given year. We consider two measures of size – book assets and sales – 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 (40) 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_0^s in Figure 2 for horizons s of 1 to 5 years. Panels A and B 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 C 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.

In sum, the empirical results in this section document that heterogeneity in firm growth opportunities is economically significantly related to the heterogeneity in the level of executive pay. These findings are identified using firm-level deviations from year (and industry) fixed effects. Given the attention that aggregate trends in executive pay have attracted in the literature, we next shift our focus on the ability of our proposed mechanism to explain the aggregate fluctuations in inequality in executive pay.

3.2 The dynamics of pay inequality

In this section, we examine more closely the observed time-series of pay inequality through the lens of our structural model. Given a set of observable variables, we explore how the time-series of pay inequality would have looked like in the data under the null of the model. Recall that in our model, the level of pay inequality – defined in equations (30) and (31) – is summarized by the ratio of the two skill prices ν and w, which itself is a monotonically increasing function of the state variable ω. Our first step is therefore to construct an analogue for ω in the data.

The model suggests five observable variables that can be mapped directly to ω. A natural choice is the value of new projects – the KPSS measure in (32) aggregated across firms, and scaled either by aggregate output or the aggregate stock market capitalization. Additionally, we use the investment-to-capital ratio and the investment-to-output ratio; in the model, both variables are increasing functions of ω. Last, we also use the cross-sectional dispersion in firm investment rates as an additional proxy for ω. In the model, increases in the value of new projects magnify the differences in match quality across firms, and hence the observed dispersion in investment rates. Panel A of Figure 3 shows that all of these variables are monotonically related to the state variable ω in simulated data.

We plot the time series of these variables in Panel B of Figure 3. We see that these series are fairly correlated with each other; the average correlation between them is 55%. However, each one individually is likely an imperfect proxy for ω. To alleviate measurement error, we use the first principal component across these five series as our empirical proxy for ω – labeled $\hat{\omega}$.[11] The estimated principal component loads positively on all five variables, and accounts for approximately 70% of their realized variance. Figure 4 compares the time series of inequality in executive pay using the Frydman-Saks data and our model-implied proxy $\hat{\omega}$.

The last step is to use this empirical measure of ω to construct the model-implied time-series of inequality. Using the estimated parameters in Section 2, we simulate one long sample of the model (10,000 years). We use the simulated data to estimate a projection of the two types of pay inequality in the model on the state variable ω (normalized to unit standard deviation). A linear projection of log pay inequality on ω works quite well (R^2 in excess of 93%), while higher order projections yield similar results. We then use these projection

---

[11]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. Probabilistic 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. Further, replicating the empirical procedure of extracting $\hat{\omega}$ as a principal component works quite well in simulated data: the correlation between $\hat{\omega}$ and the true ω is in excess of 95% in a long simulation sample of 10,000 years.

coefficients to construct model-implied measures of pay inequality in the data using the empirical proxy $\hat{\omega}$.

Figure 4 compares the time-series of the model-implied levels of pay inequality to the actual data. The solid line in Panel A of Figure 4 presents the earnings gap between executives and workers, defined as the ratio of the average total executive pay relative to average worker earnings. Consistent with the evidence in Frydman and Saks (2010), the mean executive-worker pay ratio 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.[12] 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 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. Since then, this ratio has come down somewhat, but is still higher than the 1930s.[13] The dashed line represents the time-series of pay inequality implied by the model. The correlation between the two series is 46% (HAC standard error of 16%). We see that the model can capture the sharp increase in pay inequality during the 1990s and 2000s, as well as the subsequent decline. To some extent, it can also capture the decline in the 1940s. However, the model fails to capture the flat pattern of inequality during the 1960s to 1970s – our empirical proxy for $\hat{\omega}$ implies that inequality should have increased during that period.

The solid line in Panel B of Figure 4 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-firm executive pay inequality in our sample, the standard deviation of log executive pay was equal to 1.02, more than twice its level in the 1940s. As before, the dashed line represents the level of pay inequality implied by the model. The correlation between the two series is 69% (HAC standard error of 13%). We see that the model does a particularly good job fitting the increase and subsequent decline in the dispersion in executive pay across firms during the 1990s to the mid-2010s. As before, the model can also somewhat explain the decline in pay inequality in the 1940s. However, the model again fails during the 1960s to 1970s; the model implies higher dispersion in pay across firms than what we see in the data.

---

[12]The 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.

[13]Changes 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. The data towards the end of the sample include a smaller number of firms, so they should be interpreted with caution.

The decoupling between investment opportunities (ω in the model) 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 very 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 ω 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 extremely low frequencies. Next we analyze whether the model has more predictive power for medium-run fluctuations. We do so in two ways.

First, we filter both the empirical and the model-implied measures of pay inequality using the band-pass filter (Christiano and Fitzgerald, 2003). 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 middle row (Panels C and D) of Figure 4. 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 the empirical and the model-implied measures of pay inequality are higher at medium frequencies – the correlation between the data and the model-implied measures are 71% (HAC standard error of 10%) for the inequality between executives and workers, and 69% (HAC standard error of 17%) for the inequality in executive pay across firms. We see that, once the extremely low-frequency patterns are excluded, the model performs better in the 1950s to 1970s.

Second, in the bottom row (Panels E and F) of Figure 4, we compare 5-year changes in the empirical and the model-implied measures of pay inequality. The correlation between the empirical and model-implied time-series are 41–45% (with HAC standard errors of 14–19%). As before, we see that the model performs particularly well prior to the 1950s and after the 1980s. Changes in pay inequality in the 1960s and 1970s show some qualitatively similar patterns between the data and the model, but the magnitudes are off; the model implies a higher increase in inequality during this period than the one we observe in the data.

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 pay inequality. The model performs particularly well in the recent period (after the

1990s) and to some extent prior to the 1950s. However, the model misses the decline in pay inequality during the middle part of the sample. Clearly, though fluctuations in the economy's investment opportunities were important, the model omits other factors that drove fluctuations in pay inequality during this period.

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, executives are paid more in firms with more growth opportunities, and that medium-run fluctuations in the aggregate time series in executive pay inequality are correlated with aggregate measures of growth opportunities.

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.

References

Aghion, Philippe, Ufuk Akcigit, Antonin Bergeaud, Richard Blundell, and David Hemous. "Innovation and Top Income Inequality." Working Paper 21247, National Bureau of Economic Research (2015).

Autor, David H., Lawrence F. Katz, and Alan B. Krueger. "Computing Inequality: Have Computers Changed the Labor Market?" *The Quarterly Journal of Economics*, 113(1998), 1169–1213.

Bakija, Jon, Adam Cole, and Bradley Heim. "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data." Department of Economics Working Papers 2010-22, Department of Economics, Williams College (2008).

Bazdresch, Santiago, R. Jay Kahn, and Toni M. Whited. "Estimating and Testing Dynamic Corporate Finance Models." Working paper, University of Michigan (2016).

Christiano, Lawrence J. and Terry J. Fitzgerald. "The Band Pass Filter." *International Economic Review*, 44(2003), 435–465.

Comin, Diego and Mark Gertler. "Medium-Term Business Cycles." *American Economic Review*, 96(2006), 523–551.

Davis, Morris A., Jonas D. M. Fisher, and Toni M. Whited. "Macroeconomic Implications of Agglomeration." *Econometrica*, 82(2014), 731–764.

Eisfeldt, Andrea L. and Camelia M. Kuhnen. "CEO turnover in a competitive assignment framework." *Journal of Financial Economics*, 109(2013), 351–372.

Erickson, Timothy and Toni M. Whited. "Treating Measurement Error in Tobin's q." *Review of Financial Studies*, 25(2012), 1286–1329.

Fernandes, Nuno, Miguel A. Ferreira, Pedro Matos, and Kevin J. Murphy. "Are U.S. CEOs Paid More? New International Evidence." *Review of Financial Studies*, 26(2013), 323–367.

Frydman, Carola. "Rising through the ranks. the evolution of the market for corporate executives, 1936-2003." Working paper, Boston University (2015).

Frydman, Carola and Raven Molloy. "Pay Cuts for the Boss: Executive Compensation in the 1940s." *The Journal of Economic History*, 72(2012), 225–251.

Frydman, Carola and Raven S. Molloy. "Does tax policy affect executive compensation? Evidence from postwar tax reforms." *Journal of Public Economics*, 95(2011), 1425 – 1437. Special Issue: International Seminar for Public Economics on Normative Tax Theory.

Frydman, Carola and Raven E. Saks. "Executive Compensation: A New View from a Long-Term Perspective, 19362005." *Review of Financial Studies*, 23(2010), 2099–2138.

Gabaix, X. and A. Landier. "Why has CEO pay increased so much?" *Quarterly Journal of Economics*, 123(2008), 49–100. Cited By 312.

Gentzkow, Matthew and Jesse M. Shapiro. "Measuring the Sensitivity of Parameter Estimates to Sample Statistics." Working Paper 20673, National Bureau of Economic Research (2014).

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

Griliches, Zvi. "Capital-Skill Complementarity." *The Review of Economics and Statistics*, 51(1969), 465–68.

Holmstrom, Bengt and Steven N. Kaplan. "The State of U.S. Corporate Governance: What's Right and What's Wrong?" *Journal of Applied Corporate Finance*, 15(2003), 8–20.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante. "The Effects of Technical Change on Labor Market Inequalities." In Philippe Aghion and Steven Durlauf, editors, "Handbook of Economic Growth," volume 1 of *Handbook of Economic Growth*, chapter 20. Elsevier (2005). 1275–1370.

Ingram, Beth Fisher and Bong-Soo Lee. "Simulation estimation of time-series models." *Journal of Econometrics*, 47(1991), 197–205.

Jones, Charles I. and Jihee Kim. "A Schumpeterian Model of Top Income Inequality." NBER Working Papers 20637, National Bureau of Economic Research, Inc (2014).

Jorda, Oscar. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95(2005), 161–182.

Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *Journal of Political Economy*, 87(1979), pp. 972–990.

Kaplan, Steven N. "CEO Pay and Corporate Governance in the U.S.: Perceptions, Facts, and Challenges." *Journal of Applied Corporate Finance*, 25(2013), 8–25.

Kaplan, Steven N. and Joshua Rauh. "It's the Market: The Broad-Based Rise in the Return to Top Talent." *Journal of Economic Perspectives*, 27(2013), 35–56.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. "Technological Innovation, Resource Allocation, and Growth." *Quarterly Journal of Economics*, forthcoming(2016).

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman. "Technological Innovation: Winners and Losers." Working Paper 18671, National Bureau of Economic Research (2015).

Krusell, Per, Lee E. Ohanian, Jos-Vctor Ros-Rull, and Giovanni L. Violante. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis." *Econometrica*, 68(2000), 1029–1053.

Liptser, R. S. and A. N. Shiryaev. *Statistics of Random Processes,* Vol. I: *General theory.* 2nd edition (2001).

Lustig, Hanno, Chad Syverson, and Stijn Van Nieuwerburgh. "Technological change and the growing inequality in managerial compensation." *Journal of Financial Economics*, 99(2011), 601–627.

Murphy, Kevin J. "Regulating Banking Bonuses in the European Union: A Case Study in Unintended Consequences." *European Financial Management*, 19((2013), 631–657.

Murphy, Kevin J. and Ján Zábojník. "CEO Pay and Appointments: A Market-Based Explanation for Recent Trends." *American Economic Review*, 94(2004), 192–196.

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

Nickerson, Jordan. "A Structural Estimation of the Cost of Suboptimal Matching in the CEO Labor Market." Technical report, Boston College (2014).

Papanikolaou, Dimitris. "Investment Shocks and Asset Prices." *Journal of Political Economy*, 119(2011), pp. 639–685.

Piketty, Thomas and Emmanuel Saez. "Income Inequality in the United States, 1913-1998." *The Quarterly Journal of Economics*, 118(2003), 1–41.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy*, 6(2014), 230–71.

Roweis, Sam. "EM algorithms for PCA and SPCA." *Advances in neural information processing systems*, (1998), 626–632.

Sekyu, Choi and Jos-Vctor Rios-Rull. "Understanding the Dynamics of the Labor Share: the Role of non-Competitive Factor Prices." *Annales d'Economie et de Statistique*, (2009), 251–277.

Smith, Clifford Jr. and Ross L. Watts. "The investment opportunity set and corporate financing, dividend, and compensation policies." *Journal of Financial Economics*, 32(1992), 263–292.

Solow, Robert M. "Investment and Technical Progess." In Kenneth J. Arrow, Amuel Karlin, and Patrick Suppes, editors, "Mathematical Methods in the Social Sciences," Stanford University Press, Stanford, CA (1960). 89–104.

Taylor, Lucian A. "Why Are CEOs Rarely Fired? Evidence from Structural Estimation." *The Journal of Finance*, 65(2010), 2051–2087.

―――. "{CEO} wage dynamics: Estimates from a learning model." *Journal of Financial Economics*, 108(2013), 79 – 98.

Terviö, M. "The difference that CEOs make: An assignment model approach." *American Economic Review*, 98(2008), 642–668. Cited By 68.

Tipping, Michael E. and Christopher M. Bishop. "Probabilistic Principal Component Analysis." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 61(1999), pp. 611–622.

Tables

Table 1: Summary Statistics: Firm-level Data

Variable	Mean	SD	p10	p25	p50	p75	p90
	A. Long sample (1936-2014)						
Exec. Compensation (1982 $m)	2.09	2.44	0.59	0.78	1.18	2.20	4.78
Compensation to Assets (%)	0.02	0.02	0.00	0.01	0.01	0.02	0.05
Book assets (log 1982 USDm)	9.31	1.28	7.75	8.47	9.23	10.13	10.93
Net Income to Assets	0.06	0.05	0.01	0.03	0.06	0.08	0.11
Stock Returns	0.10	0.27	-0.25	-0.06	0.12	0.26	0.42
Firm Innovation, $\hat{\nu}/B$ (%)	9.14	15.82	0.00	0.00	2.33	12.14	25.74
	B. Large Panel (1992-2014)						
Exec. Compensation (1982 $m)	2.45	2.91	0.53	0.85	1.53	2.93	5.20
Compensation to Assets (%)	0.24	0.40	0.02	0.05	0.12	0.26	0.55
Book assets (log 1982 USDm)	7.34	1.58	5.38	6.18	7.21	8.39	9.56
Net Income to Assets	0.04	0.12	-0.05	0.02	0.05	0.09	0.13
Stock Returns	0.19	0.59	-0.38	-0.13	0.11	0.38	0.74
Firm Innovation, $\hat{\nu}/B$ (%)	8.99	28.67	0.00	0.00	0.00	4.87	25.04
Tobin's Q (log)	1.39	1.08	0.05	0.53	1.28	2.09	2.87
Investment Rate	0.11	0.20	-0.04	0.02	0.07	0.16	0.31

Executive compensation in firm f and year t is the mean value of total pay of the top-3 managers (in the Frydman and Saks (2010) sample) and top-5 managers (in Execucomp). The compensation data in the long sample is from Frydman and Saks (2010) (extended to 2014 using Execucomp) and includes the ex-ante value of options. Compensation data from Compustat/Execucomp is TDC1. The value of book assets (Compustat: at) and compensation are deflated to year 1982 dollars using the CPI. The firm-level innovation measure $\hat{\nu}$ is from Kogan et al. (2016). Tobin's Q is book assets (at) plus market value of equity (prcc_f times csho) minus common equity (ceq) and deferred taxes (txdb) divided by property, plant and equipment (ppegt). Industry definitions are based on the SIC2 level. Investment is CAPX (Compustat: capx) scaled by PPE (Compustat: ppegt). Stock returns are obtained from CRSP.

Table 2: Model: Goodness of Fit

Statistic	Data	Model			$|D-M|$
		Mean	(5%)	(95%)	(p-value)
Aggregate quantities and prices					
Consumption growth, mean	0.015	0.013	-0.004	0.025	0.142
Consumption growth, volatility (annual)	0.038	0.041	0.034	0.051	0.095
Consumption growth, volatility (long-run)	0.042	0.056	0.035	0.081	0.091
Investment-to-output ratio, mean	0.094	0.092	0.043	0.137	0.646
Investment-to-output ratio (log), volatility	0.305	0.311	0.147	0.535	0.875
Investment growth, volatility	0.140	0.119	0.085	0.140	0.190
Investment and consumption growth, correlation	0.467	0.339	0.163	0.518	0.579
Net payout to assets, coefficient of variation	0.575	0.441	0.170	0.941	0.089
Risk-free rate, mean	0.015	0.014	0.005	0.027	0.786
Executive pay and inequality					
CEO tenure (years), median	6.260	6.317	4.212	7.121	0.393
Elasticity of executive pay to firm size	0.386	0.394	0.000	0.757	0.855
Dispersion in log executive pay, TS mean	0.545	0.597	0.397	0.747	0.659
Dispersion in log executive pay, log, TS volatility	0.278	0.249	0.114	0.398	0.862
Mean executive pay to worker, log, TS mean	3.753	3.754	3.206	4.143	0.985
Mean executive pay to worker, log, TS volatility	0.521	0.243	0.132	0.386	0.000
Cross-sectional (firm) moments					
Firm investment rate, IQR	0.171	0.216	0.179	0.283	0.001
Firm investment rate, serial correlation	0.203	0.111	0.001	0.231	0.002
Firm Innovation (KPSS), IQR	0.519	0.608	0.374	0.822	0.121
Firm Innovation (KPSS), serial correlation	0.251	0.148	0.105	0.268	0.311
Firm profitability, IQR	1.001	0.976	0.849	1.216	0.309
Firm profitability, serial correlation	0.742	0.746	0.721	0.760	0.652
Wald Distance criterion (p-value)	0.000				

This table reports the fit of the model to the statistics of the data that we target. All statistics are reported at annual frequencies. See the main text and the Online Appendix for more details on the definition of these statistics and on the estimation and simulation of the model. In addition to presenting the mean, 5th and 95th percentiles of the estimates from the model, in the last column we also report the p-values of a t-test to asses the statistical significance of the differences between the model and the data. Standard errors take into account both sampling and simulation uncertainty. The last row reports the p-value associated with an overall measure of model fit, that is, that the minimized value in (34) is statistically different from zero.

Table 3: Parameter Estimates

Parameter	Symbol	Estimate	SE
Executive Labor Market			
Executive share of match surplus	η	0.3835	0.1531
Worker effective supply of labor	h	0.0022	0.0006
Executive effective supply of labor	e	0.0165	0.0043
Proportional termination cost	c	0.7579	0.1913
Unconditional match quality	$\bar{p}$	0.0669	0.0140
Involuntary turnover rate	β	0.0103	0.0052
Technology, Preferences and Production			
Rate of time preference	ρ	0.0047	0.0043
Decreasing returns to investment	α	0.3533	0.0381
Depreciation rate	δ	0.0309	0.0251
Disembodied technology growth, mean	μ_x	0.0154	0.0193
Disembodied technology growth, volatility	σ_x	0.0805	0.0109
Embodied technology growth, mean	μ_ξ	0.0005	0.0280
Embodied technology growth, volatility	σ_ξ	0.1619	0.0078
Project mean arrival rate, low quality match	λ_L	0.0534	0.0245
Project mean arrival rate, difference high vs low quality match	λ_D	0.6889	0.3324
Project-specific productivity, long-run volatility	v_u	1.6803	1.1089
Project-specific productivity, persistence	κ_u	0.3325	0.0187

This table reports the estimated parameters of the model. See the main text and the Online Appendix for details on the estimation of the model.

Table 4: Executive Pay and Innovation: KPSS Measure of Growth Opportunities

$\log X_{f,t}$	(1)	(2)	(3)	(4)	(5)
	A. Long sample (1936–2014)				
$\hat{\nu}_{f,t}/B_{f,t}$	0.142***	0.092***	0.051***	0.052***	0.045**
	(0.025)	(0.019)	(0.018)	(0.018)	(0.021)
$\log B_{f,t}$		0.308***	0.308***	0.309***	0.314***
		(0.019)	(0.020)	(0.020)	(0.027)
$\log(1 + ROA_{f,t})$			2.431***	2.223***	2.087***
			(0.304)	(0.306)	(0.262)
$\log(1 + R_{f,t})$				0.150***	0.151***
				(0.033)	(0.031)
Observations	5299	5299	5299	5299	5299
R^2	0.906	0.935	0.939	0.939	0.952
	B. Large panel (1992–2014)				
$\hat{\nu}_{f,t}/B_{f,t}$	0.236***	0.107***	0.105***	0.103***	0.080***
	(0.051)	(0.015)	(0.016)	(0.015)	(0.011)
$\log B_{f,t}$		0.406***	0.404***	0.408***	0.360***
		(0.009)	(0.009)	(0.009)	(0.014)
$\log(1 + ROA_{f,t})$			0.250***	0.172**	0.433***
			(0.066)	(0.069)	(0.067)
$\log(1 + R_{f,t})$				0.105***	0.066***
				(0.022)	(0.018)
Observations	33551	33551	33551	33551	33551
R^2	0.261	0.627	0.629	0.632	0.785
Year FE	Y	Y	Y	Y	Y
Industry (SIC2) FE	Y	Y	Y	Y	-
Firm FE	-	-	-	-	Y

This table reports estimates of equation (39) in the text. The dependent variable $X_{f,t}$ is the firm level of executive compensation, defined as the average compensation (including the ex-ante value of options) in a given year of the firm's top-3 executives (long sample, i.e. Frydman-Saks) or the firm's top 5 executives (large panel, i.e. Execucomp). We measure ν by the firm-level innovation measure from Kogan et al. (2016) (scaled by the book value of assets), firm size B by the book value of assets, firm profitability ROA by the ratio of net income scaled by book assets, and R_t by the firm's stock return. Depending on the specification, we include industry (SIC2) or firm fixed effects. Standard errors are clustered by firm and year, and are presented in parentheses.***,** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 5: Executive Pay and Innovation: Other Measures of Growth Opportunities (Large Panel)

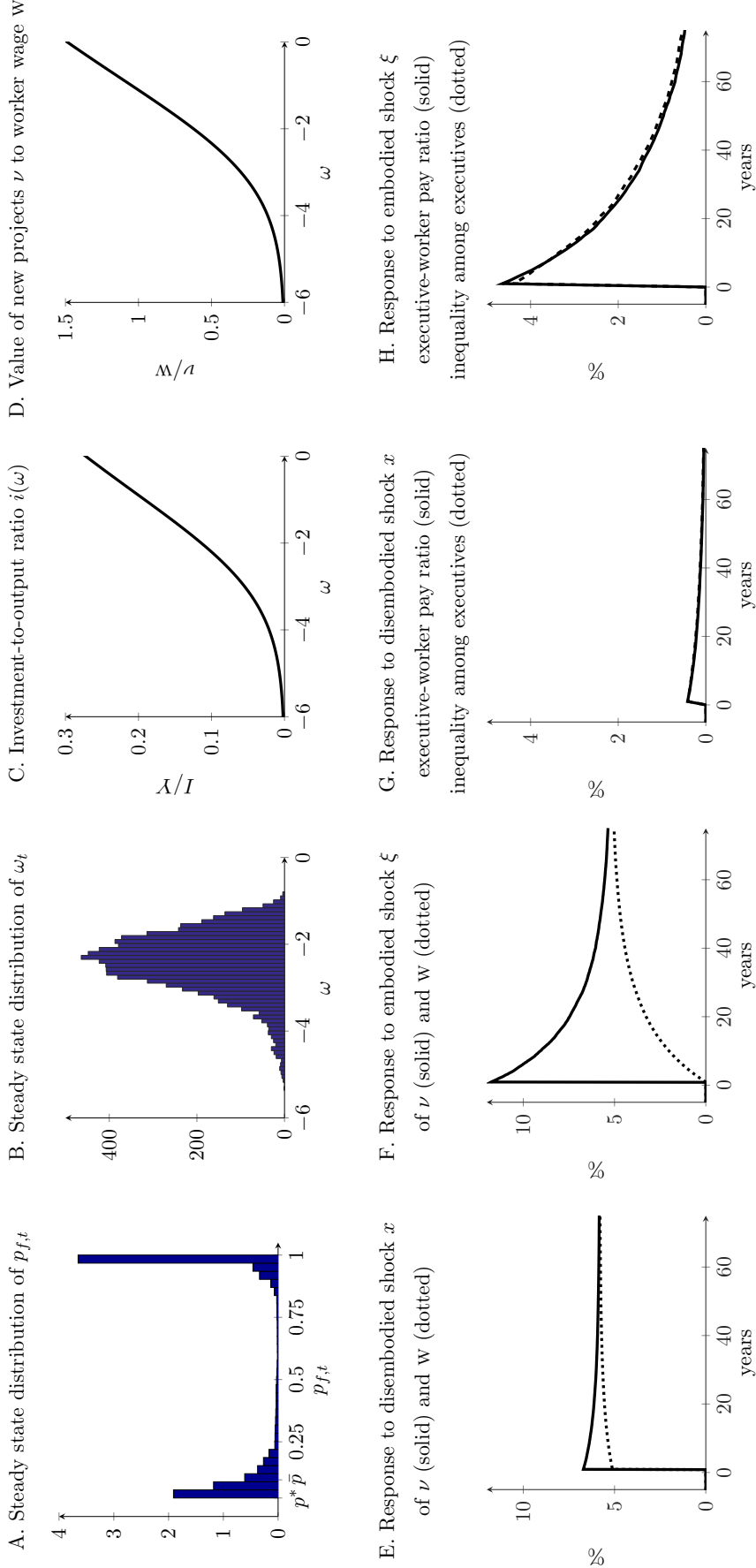| $\log X_{f,t}$	(1)	(2)	(3)	(4)	(5)
	A. Tobin's Q				
$\log Q_{f,t}$	0.165***	0.234***	0.236***	0.229***	0.259***
	(0.025)	(0.018)	(0.019)	(0.019)	(0.025)
R^2	0.250	0.670	0.671	0.672	0.806
	B. Firm Investment				
$i_{f,t}$	0.051***	0.134***	0.131***	0.130***	0.083***
	(0.017)	(0.010)	(0.010)	(0.011)	(0.011)
R^2	0.207	0.633	0.634	0.638	0.785
	C. Index of Growth Opportunities (PCA)				
$G_{f,t}$	0.201***	0.225***	0.224***	0.219***	0.182***
	(0.025)	(0.010)	(0.010)	(0.010)	(0.012)
R^2	0.237	0.656	0.657	0.659	0.792
Observations	33551	33551	33551	33551	33551
Industry FE	Y	Y	Y	Y	-
Year FE	Y	Y	Y	Y	Y
Firm FE	-	-	-	-	Y
Size (book assets)	-	Y	Y	Y	Y
ROA	-	-	Y	Y	Y
Stock Return	-	-	-	Y	Y

This table reports estimates of equation (39) in the text using three alternative measures of growth opportunities: (A) Tobin's Q, defined as book assets (at) plus market value of equity (prcc_f times csho) minus common equity (ceq) and deferred taxes (txdb), divided by property, plant and equipment (ppegt); (B) the firm investment rate, defined as the ratio of capital expenditures (capx) to plant and equipment (ppegt); and (C) the first principal component across $\log Q$, investment rate and $\hat{\nu}/B$. See the notes to Table 4 for additional variable definitions. Depending on the specification, we include industry (SIC2) or firm fixed effects. Standard errors are clustered by firm and year. ***,** and * indicate significance at the 1%, 5% and 10% level, respectively.

Table 6: Model: Executive Pay and Growth Opportunities

| $\log X_{f,t}$	(1)	(2)	(3)	(4)	(5)
	A. Firm Innovation (KPSS)				
$\hat{\nu}_{f,t}/B_{f,t}$	0.194	0.161	0.161	0.178	0.103
	(0.031)	(0.021)	(0.021)	(0.023)	(0.021)
R^2	0.119	0.895	0.904	0.915	0.955
	B. Tobin's Q				
$\log Q_{f,t}$	-0.188	1.345	1.368	1.415	1.463
	(0.218)	(0.245)	(0.246)	(0.255)	(0.252)
R^2	0.119	0.895	0.904	0.915	0.955
	C. Firm Investment				
$i_{f,t}$	0.206	0.178	0.178	0.210	0.097
	(0.045)	(0.044)	(0.044)	(0.050)	(0.021)
R^2	0.110	0.493	0.493	0.498	0.805
	D. Index of Growth Opportunities (PCA)				
$G_{f,t}$	0.212	0.239	0.239	0.293	0.165
	(0.055)	(0.065)	(0.065)	(0.071)	(0.024)
R^2	0.110	0.524	0.525	0.535	0.811
Observations	30000	30000	30000	30000	30000
Year FE	Y	Y	Y	Y	Y
Firm FE	-	-	-	-	Y
Size	-	Y	Y	Y	Y
ROA	-	-	Y	Y	Y
Stock Return	-	-	-	Y	Y

This table reports estimates of equation (39) in simulated data from the model using four alternative measures of growth opportunities: (A) the value of new projects (constructed using equation (33) in the text) scaled by the book value of capital; (B) Tobin's Q, defined as the ratio of the market value of the firm to the book value of capital; (C) firm investment rate, defined as the change in log capital; and (D) the first principal component across (A), (B) and (C). We compute model standard errors as the standard deviation of parameter estimates across $S = 100$ simulations. See the main text and the Online Appendix for additional details on the model estimation and simulation methodology.

Figure 1: Forces determining pay inequality in the model

A. Steady state distribution of $p_{f,t}$

B. Steady state distribution of ω_t

C. Investment-to-output ratio $i(\omega)$

D. Value of new projects ν to worker wage w

E. Response to disembodied shock x of ν (solid) and w (dotted)

F. Response to embodied shock ξ of ν (solid) and w (dotted)

G. Response to disembodied shock x executive-worker pay ratio (solid) inequality among executives (dotted)

H. Response to embodied shock ξ executive-worker pay ratio (solid) inequality among executives (dotted)

Panels A and B of the figure plot the steady-state distribution of perceived match quality $p_{f,t}$ and the state variable ω_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. Panels C and D plot the investment-to-output ratio and the ratio of the value of investment opportunities ν_t to the equilibrium worker wage w_t. Panels E through H plot the impulse response of the value of investment opportunities ν_t, the equilibrium worker wage w_t, inequality between executives and workers (defined in equation (30), and inequality among executives (defined in equation (30)), and inequality among executives in our model. Impulse responses are constructed by taking into account the nonlinear nature of equilibrium dynamics: we introduce an additional one-standard deviation shock at time $t = 0$ without altering the realizations of all future shocks. The impulse responses are computed at the mean of the stationary distribution of ω. We report the log difference between the mean response of the perturbed and unperturbed series (multiplied by 100).

47

Figure 2: Executive pay and future firm growth

A. Long sample (1936-2014)


B. Large panel (1992-2014)
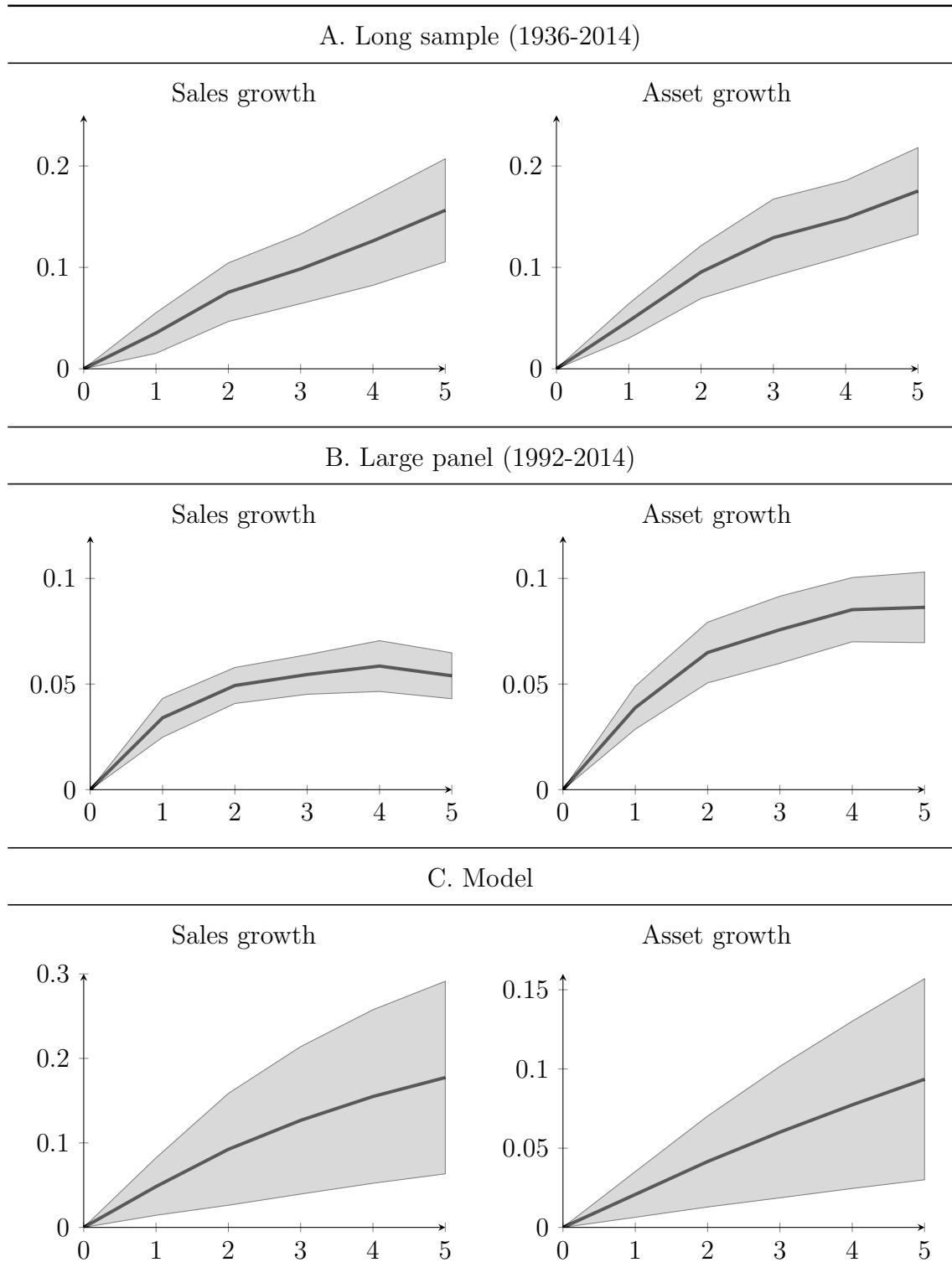
C. Model

This Figure presents estimates of equation (40) in the long sample (panel A), the large panel (panel B) and in simulated data from the model (panel C). Shaded areas correspond to 90% confidence intervals. In Panels A and B (data), standard errors are clustered by firm and year. In Panel C (model), standard errors are computed as the standard deviation of parameter estimates across simulations.
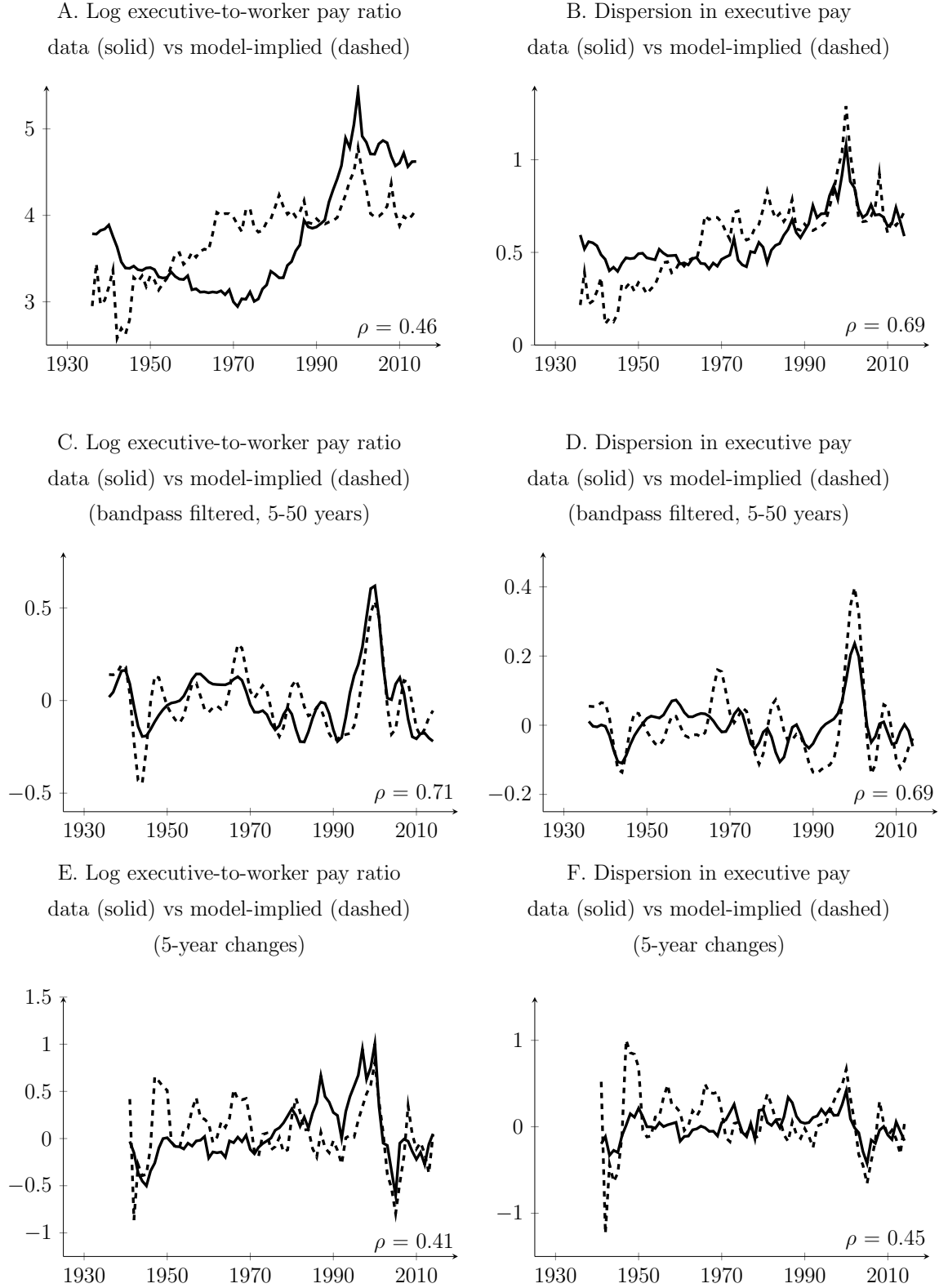
Figure 3: Model-implied proxies for ω
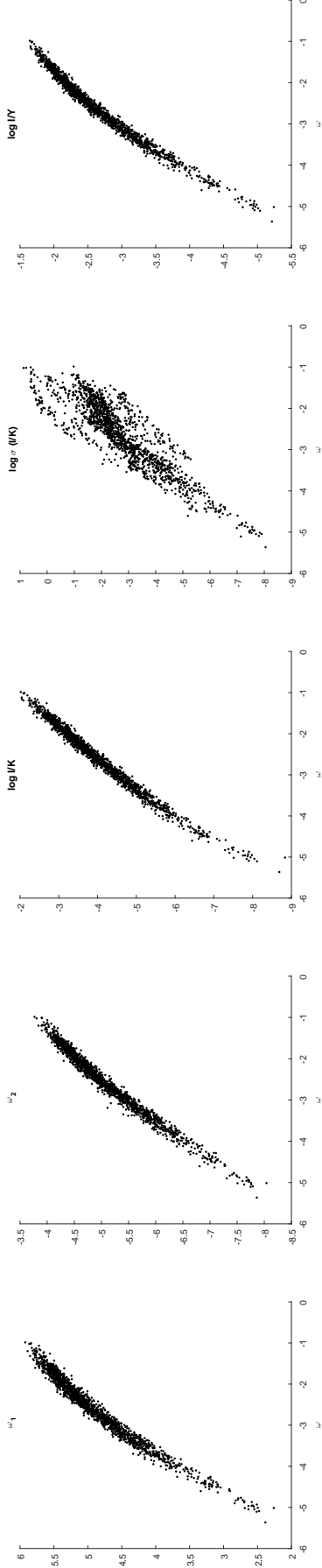
A. Model

B. Data
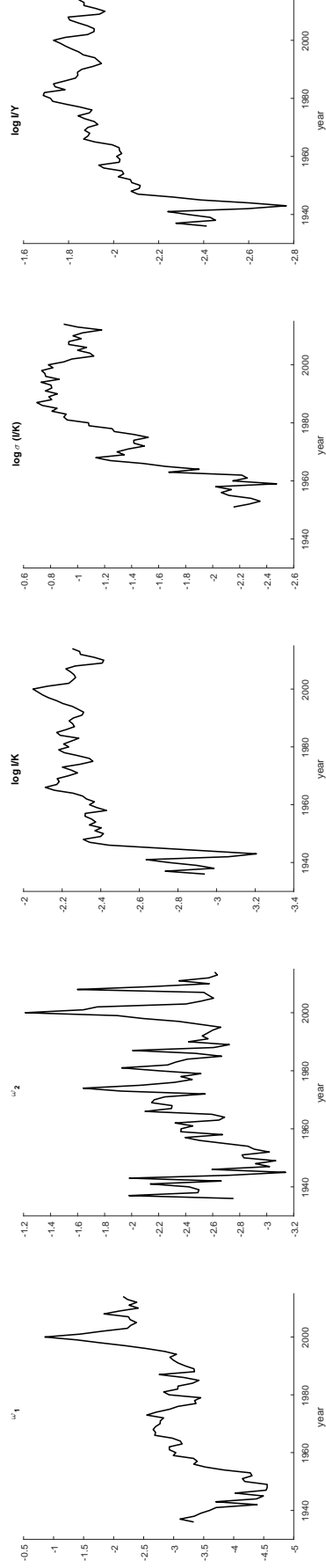
This figure examines five observable variables that can be mapped directly to ω in our model: the logarithm of the value of new projects – the KPSS measure in (32) aggregated across firms, and scaled either by aggregate output (ω_2) or the aggregate stock market capitalization ($\hat{\omega}_2$); the logarithm of the investment-to-capital ratio IK_t and the investment-to-output ratio IY_t; and the cross-sectional dispersion in firm investment rates, $\sigma_t(i_{f,t})$. Panel A shows that these five variables are monotonically related to the state variable ω. Panel B plots the time series for these variables in the data. The variables ω_1 and ω_2 are constructed using the innovation measure in Kogan et al. (2016); aggregate market capitalization is obtained from CRSP. GDP is the nominal output from the BEA (Table 1.1.5, row 1). The investment to output ratio is constructed as the ratio of non-residential private investment (Table 1.1.5, row 9) divided by GDP (Table 1.1.5, row 1). The investment-to-capital ratio is non-residential private investment (Table 1.1.5, row 9) divided by non-residential fixed assets (Table 1.1, row 4). The cross-sectional dispersion in firm investment rates is estimated using Compustat.

Figure 4: The Dynamics of Pay Inequality, 1936 to 2014

A. Log executive-to-worker pay ratio
data (solid) vs model-implied (dashed)

$\rho = 0.46$

B. Dispersion in executive pay
data (solid) vs model-implied (dashed)

$\rho = 0.69$

C. Log executive-to-worker pay ratio
data (solid) vs model-implied (dashed)
(bandpass filtered, 5-50 years)

$\rho = 0.71$

D. Dispersion in executive pay
data (solid) vs model-implied (dashed)
(bandpass filtered, 5-50 years)

$\rho = 0.69$

E. Log executive-to-worker pay ratio
data (solid) vs model-implied (dashed)
(5-year changes)

$\rho = 0.41$

F. Dispersion in executive pay
data (solid) vs model-implied (dashed)
(5-year changes)

$\rho = 0.45$

This figure compares the relation between pay inequality in the data (solid lines) and the level of pay inequality that is implied by the model given a set of observables (dashed line). The set of observable variables that are used to construct the model-implied time-series of pay inequality are plotted in Panel B of Figure 3. The top row of the figure (Panels A and B) plots these series in levels; the middle row (Panels C and D) presents the band-pass filtered series, keeping frequencies of 5 to 50 years; the last row (Panels E and F) plots 5 year changes for each series, defined as $\Delta x_t = x_t - x_{t-5}$.