The Ratchet Effect of Transparency in Executive Compensation

José Guedes

October, 2018

Abstract

How does the mandated disclosure of executive compensation affect the dynamics of compensation within a peer group of firms? We argue in this paper that, under plausible conditions, it will trigger a ratchet effect in the mean level of executive pay within the peer group accompanied, hand in hand, by a sustained shrinkage in the variance of executive pay. The mean and the variance of executive compensation change periodically until they reach their new steady state values under the regime of full disclosure, a higher one in the case of the mean and a lower one in the case of the variance; they change quickly immediately following the kick off of the full disclosure regime but their increments taper off gradually as they approach their new steady values.

The duration of convergence process of the mean pay depends critically on how firms and shareholders develop forecasts of the future wages paid to executives within the peer group. When participants are rational and fully understand the equilibrium implications of the wage setting process at individual firms on the distribution of aggregate wages within the peer group, the convergence occurs quickly. However, if participants are less than fully rational or less than perfectly well-informed, and form their forecasts based on simple updating rules applied to disclosed compensation data, the ratchet effect in the mean wage can last for a long time.

---

1 Católica-Lisbon School of Business and Economics, Universidade Católica Portuguesa. jcg@ucp.pt
I. Introduction

Edmands, Gabaix and Jenter (2017) document a steady rise in CEO pay between the early 1970s and the mid 2000s. In particular, they report a six-fold increase between 1980 and 2014, with annual growth rates in excess of 10% during the 1990s. A parallel trend towards greater disclosure of CEO pay of listed firms has taken place over the past decades. Since the 1934 Securities Exchange Act, regulators have engaged in a process of ever increasing disclosure requirements that has continued unabated until the recent past. In section II of this paper we review the major regulatory events that increased or enhanced the level and scope of disclosure and see that their timing maps well into periods of quick CEO pay growth.

The coincident timing these two trends prompts us to ask the question: Can mandated disclosure cause an escalation in CEO pay? Is it possible that simply giving firms access to information about pay conditions of CEOs within their peer group leads them to pay more to their own CEOs? We respond affirmatively to these questions, by uncovering a mechanism through which the observation of the wages paid to CEOs at similar firms enables each individual CEO to strike better wage deal with her own firm.

The paper develops a model that explores the consequences of transparency in CEO wages on compensation dynamics within a peer group of firms (henceforth, referred to as the industry). Our major finding is that when participants (i.e., managers and shareholders) face uncertainty about the exact value of a structural parameter underlying the wage setting process at each and every individual firm in the industry, the disclosure of wages allows them to learn about such parameter, thereby reducing the uncertainty about the value of their outside options in case of a negotiation break-down. Although risk-neutral shareholders are unaffected by such reduction in uncertainty, risk-averse managers experience an increase in the expected utility of their outside option, which empowers them to strike a better wage deal with shareholders.

We examine a setup where participants, prior to the disclosure of executive pay, hold disparate estimates of an unknown parameter that influences the distribution of wages and, once wages start to be publicly disclosed, gradually converge their estimates toward the true value of the parameter. The wage generated periodically by each individual firm depends on the forecasts of the mean and the variance of the industry wages held by the firm’s manager and shareholders in the preceding period, so realized wages depend on forecasted wages. Our benchmark is the case in which participants have rational expectations, i.e., the case in which participants build forecasts that, if their beliefs about the value of the unknown parameter are correct, yield a distribution of realized wages that matches their forecasted distribution. This rational expectations benchmark, however, is quite demanding for it requires participants to be sufficiently smart and well-informed to figure out the implications of the wage setting process at individual firms on the equilibrium distribution of industry wages. We thus explore an alternative in which managers and shareholders are less sophisticated and forecast the mean and the variance of future industry wages by applying a simple weighing scheme to the historical data on disclosed wages. In this
alternative, which we refer to as adaptive expectations, participants simply adjust, on a periodic basis, their forecasts to reflect the most recently disclosed wage data.

Under both rational expectations and adaptive expectations we find that mandated disclosure of executive pay triggers a ratchet process in the mean industry wage that pulls it upward in successive increments until the initially disparate forecasts of the mean industry wage held by participants have fully converged to the new steady state value of the mean wage, i.e., to the mean wage that is consistent with the full resolution of uncertainty about the exact value of the unknown structural parameter. Initially, the mean wage rises rapidly but, as time passes, the increments taper off since additional batches of disclosed wages reveal less and less information. The variance of industry wages, on the other hand, undergoes a parallel downward process, reflecting the increasing homogeneity of the forecasts among participants. The rate of change of the variance displays the same time pattern as that of the mean industry wage, initially changing at a high clip but slowing down with the passage of time.

The duration of the adjustment process, however, is quite different under rational and adaptive expectations. Under the former, the mean industry wage converges quickly to its new steady value following the adoption of the mandated disclosure regime; under the latter it takes much longer to converge. We run simulations that show that under rational expectations 95% of the adjustment in the mean industry wage is completed within the first two periods following mandated disclosure whereas, under adaptive expectations, it takes over seven periods to achieve the same level of adjustment. Hence, deviations from rational expectations create a long-lasting ratchet effect in the mean wage of industry executives.

In our model managerial effort plays no role and, therefore, the wage paid to executives consists solely of fixed pay. We abstract from considerations about incentive effects and moral hazard and focus on the participation constraints of managers and shareholders. At each individual firm the wage of the manager is set according to a fixed proportional rule that splits the surplus created by the relationship between the manager and shareholders. All the action in the wages at individual firms – and consequently, in the distribution of industry wages - comes from shifts in the participation constraints occurring as participants revise their forecasts of the mean and variance of industry wages to reflect the information revealed by newly disclosed wages. As in Oyer (2004), the model is developed to emphasize the role played by the participation constraint channel on the dynamic of wages.

The rest of the paper is organized s follows. Section II summarizes the relevant literature; Section III presents the model; finally, section IV concludes.
II. Literature review

Edmans, Gabaix and Jenter (2017) provide a summary of the history of the regulation of executive compensation’s disclosure in the US. There are four critical moments in such history. The first one, arguably the most important, is the adoption in 1934 of Securities Exchange Act. It established the legal and regulatory framework for mandated pay disclosure of listed companies (in particular, it required the three highest-paid executives to disclose their compensation in the proxy statement).

A second key moment occurred in 1978, when disclosure rules were extended to the top five executives in listed companies and the information submitted by executives was expanded by requiring them to fill out a Summary Compensation Table (SCT).

A third moment took place in 1992, requiring the SCT to include more extensive, detailed and standardized information, namely sub-tables with information about the number of newly awarded options and stock appreciation rights (SARs), the number of option and SARs held at the end-of-the-year, the options and SARs exercised during the year and long-term incentive plans.

In 2006, a fourth relevant moment took place. New disclosure rules required the SCT to contain the value of new option grants and mandated the inclusion of a new table – the Pensions Benefit Table – with information about the present value of accumulated pension benefits, plus payments during the current year. As a result of the 2006 disclosure reform, external parties were able to compute a total compensation number for the first time. The 2006 reform, additionally, required firms to include a Compensation Discussion and Analysis (CD&A) section, a narrative description of the determinants and objectives of their compensation policies.

Several empirical studies have examined how these regulatory events affected executive compensation. Their common finding is that regulatory innovations that mandated the disclosure of executive compensation or increased the scope and level of detail of disclosed information led to a boost in the mean level of executive pay.

Mas (2016) studies the impact on CEO compensation resulting from the 1934 Securities Exchange Act. He finds an increase in average CEO pay; a compression in the wage distribution of CEOs, driven largely by a sharp reduction in the variance of residual compensation, i.e., compensation left unexplained by market cap and 2-digit SIC industry dummies; a reduction in the pay-for-performance sensitivity; an increase in compensation within the lowest percentiles of the conditional wage distribution of CEOs and no significantly change in top percentiles; and a precipitous decline at the very top of the wage distribution (above the 98th percentile).

Perry and Zenner (2001) show that the tax legislation in 1992 that capped the corporate income tax deduction of non-performance-related compensation at one million dollars (section 162(m) of the Internal Revenue Code) and the 1992 SEC rule requiring enhanced disclosure on executive compensation, were followed by a dramatic increase in real compensation levels, with all the components of compensation going up. Only salaries close or above the one million dollar
benchmark experienced a slowdown in their growth rate. The authors also find that most firms that reduced salaries to a level at or below one million dollars did so in response to 162(m), but offset the salary reductions with an increase in other components of executive pay.

Faulkender and Yang (2012) study the effects of 2006 regulation on the disclosure of peer group members used for the benchmarking of compensation. They find that firms did not select less biased peers after enhanced disclosure. If anything, the bias to select highly paid peers became stronger.

Gipper (2016) examines the impact on executive compensation from the 2006 reform. He finds that average compensation levels went up, with pay increases concentrated among managers with shorter pay histories, at smaller firms and in industries with higher variance of pay. He also finds that pay setting became more rigid (i.e., more rules based). Finally, in contrast to Mas (2016), he reports that pay dispersion increased.

Evidence outside the US also supports the view that the mandatory disclosure of executive compensation (or the mandatory adoption of enhanced disclosure requirements) raises compensation levels.

Park, Nelson and Huson (2001) analyze how the Ontario Securities Commission’s 1993 rule mandating Canadian firms trading in the Toronto Stock Exchange to disclose the amount and composition of the individual compensation of the 5 highest paid executives (before 1993 firms were only required to disclose the aggregate total compensation of all executives), affected executive compensation. They report that such rule led to an increase in total executive pay in real terms, to an increased reliance on incentive components of pay and to an increase in the pay-for-performance sensitivity.

Evidence from Germany corroborates these findings. Schmidt (2012) exploits the staggered adoption of a new corporate governance rule passed in 2006 in Germany requiring listed firms to disclose their executive compensation on an individual basis. He finds that the average compensation of executives who were previously less informed about the compensation paid externally at similar firms went up, compared to the compensation earned by their colleagues in the board who were better informed about the value of their outside option.

Also relevant to our paper, are empirical studies examining the role of peer benchmarking in the determination of CEO pay. Bizjak, Lemmon and Naveen (2008) report evidence of competitive benchmarking of CEO pay against peers, with such benchmarking strongly influencing the level of CEO compensation. Their results suggest that such practice is used to gauge the market wage necessary to retain valuable human capital. According to the authors, the findings of Garvey and Milbourn (2006) that CEOs are paid more for good luck than they are punished for bad luck, is not due the CEO entrenchment and rent seeking (as argued, for example, by Bertrand and Mullainathan (2001) and Bebchuck and Fried (2004)), but is explained by the firm’s desire to adjust pay in consonance with variations in the value of CEO’s outside option. These results are
consistent with Oyer (2004) who posits that the CEO’s outside opportunities are more attractive in broader market upswings.

O Reilly, Main and Crystal (1988) analyze the drivers of the cash component of executive compensation (salary plus bonuses) using a survey of 105 listed firms distributed across nine industries. Besides strong industry-specific effects, he finds a strong and consistent influence of social comparisons (or peer benchmarking) on CEO pay, which they capture by the salary levels of compensation committee members or other outside members of the board of directors who are executives in other firms (an increment of 100,000 USD per annum in the average salary of the outside directors on the compensation committee, the expected salary of the focal CEO will rise by 51,000 USD per annum).

As far as theory is concerned, a few papers model the effects of enhanced corporate disclosure on executive pay outcomes, but none takes a perspective similar to ours.

Acharya and Volpin (2010) develop a model in which managerial incentive contracts are a substitute for strong corporate governance. When a firm decides to go for a mix of weak governance and high incentive pay it raises the value of the outside option of managers at other firms, thereby biasing the trade-off between incentive pay and strong governance faced by these firms toward high incentive pay and weak governance. Due to this externality, the overall level of governance in the economy is excessively low whereas the level of incentive pay is excessively high when corporate governance arrangements made by firms (including the level and structure of managerial compensation) are publicly observable.

In Hermalin and Weisbach (2012), enhanced disclosure of managerial pay and performance improves the monitoring of managers by shareholders, boosting firms’ profits (e.g., by firing the manager following poor performance) while reducing the informational rents earned by managers. Hence, enhanced disclosure requirements should lead to higher managerial compensation. The increase in compensation results from (i) an increase in the reservation utility of the manager and (ii) the ability of the manager to capture part of the incremental profits through bargaining.

Hayes and Schaefer (2009) model the so-called Lake Wobegon effect. Its premise is that no firm wants to admit to having a below average CEO and so avoid giving its CEOs a pay package that lags market expectation. The upshot is a collective effect characterized by a spiral in executive pay. In their model, firms observe a private signal about the productivity of their own manager; firms maximize a weighted sum of short and long run firm value, so they may want to distort the publically observed CEO wage to affect the market’s short term-beliefs about the quality of the manager. The authors find that their model can generate a Wobegon effect only under stringent parameter restrictions, which makes them skeptical about the plausibility of such an effect.
III. The Model

Consider an industry composed of a large number of firms, each employing one manager. The skills of managers are entirely industry-specific, so their value outside the industry is zero. Firms, on the other hand, are able to operate only if run by a manager with industry-specific skills.

Firm-manager pairs generate periodically an identical amount of profit, before deducting managerial compensation, equal to $\Pi$. The profits are exogenous and do not depend on the effort of managers. We make this simplifying assumption to focus on the role of the participation constraint as a driver of the dynamics of industry wages.

Some incumbent firms die every period (forcing their managers to seek employment at other industry firms) and some new firms enter the industry (seeking to hire a manager with industry-specific skills); concomitantly, some incumbent managers retire every period (forcing their firms to find a replacement) and some new managers with industry-specific skills enter the labor market (seeking employment at existing firms). Such turnover among firms and managers creates an active managerial labor market. We assume that to achieve a match in the labor market managers and firms bear a searching cost. Specifically, a manager going to the labor market receives a wage, net of searching costs, equal to $(1-\lambda^M)w$ whereas a firm earns a profit, net of wage and searching costs, equal to $(1-\lambda^F)(\Pi-w)$, where $w$ is the manager’s wage. For the sake of simplicity, we assume that $\lambda^M$ and $\lambda^F$ are exogenous and constant. Searching costs are determined by factors such as the relative numbers of managers and firms participating in the labor market, the efficiency of the matching technology and the turnover of managers and firms in the labor market. Here, we abstract from these considerations and simply assume that these costs correspond to an exogenously fixed proportion of the benefit obtained from achieving a match.

The wage of managers of newly matched pairs is set in the same manner as that of pre-existing pairs, through a bargaining process over the joint surplus created by the relationship. Once matched, firms and managers always strike a bargain and thus do not return to the labor market in the current period.

Managers feature an identical degree of absolute risk aversion (set to be equal to one, without loss of generality) and are equally skilled; firms are risk neutral.

II.1. Wage negotiation

Each pair (manager $i$, firm $j$), either newly formed or pre-existing, negotiates independently, period by period, the wage of the manager using a Nash bargaining model. Let the fraction of the surplus of the relationship captured by the manager of pair $(i,j)$ be denoted $\eta_{ij}$. Although managers and firms know the parameter coefficient of their own pair, they do not know the parameter coefficients of other possible pairs. Let the mean and the variance of $\eta_{ij}$ in the population of all possible pairs be denoted, respectively, $\mu_{\eta}$ and $\sigma_{\eta}^2$. We assume that pairs know the variance of the
distribution, $\sigma^2_{\eta, t}$, but do not know the mean, $\mu_\eta$. Pair $(i, j)$ holds a prior distribution of $\mu_\eta$ at date $t$ that has a mean equal to $\mu^i_{\eta, t}$ and a variance equal to $\sigma^2_{\mu, t}$. We are thus assuming that pairs hold heterogeneous beliefs about the mean of the distribution but common beliefs about the variance.

The reservation utility of the manager is determined by how much she can get if she abandons the relationship and seeks a new employer in the labor market. The risk-free equivalent outside option of manager $i$ at date $t$ is written as

$$w^\text{min}_{ij,t} = (1 - \lambda^M)\mu^M_{ij,t} - \left(1 - \lambda^M\right)^2 \sigma^2_{\text{predictive}}$$

where $\mu^M_{ij,t}$ and $\sigma^2_{\text{predictive}}$ are, respectively, the mean and the variance of the predictive distribution, held by manager $i$ of pair $(i, j)$ at date $t$, of the wage earned by employed managers in period $t$; $\mu^M_{ij,t}$ represents the wage that the manager expects to obtain if she goes to the labor market while $\sigma^2_{\text{predictive}}$ represents the uncertainty about such wage that is perceived by the manager.

Since firms are risk-neutral, their reservation utility is determined by how much they expect to obtain if they attempt to replace their manager with a new one hired in the labor market. If firm $j$ goes to the labor market to hire a manager its expected profit net of labor costs is

$$(1 - \lambda^F) (\Pi - \mu^F_{j,t})$$

where $\mu^F_{j,t}$ is the mean of the predictive distribution, held by firm $j$ at date $t$, of the wage paid to employed managers in period $t$. Hence, the wage of the manager from pair $(i, j)$ cannot exceed

$$w^\text{max}_{ij,t} = \mu^F_{j,t} + \lambda^F (\Pi - \mu^F_{j,t})$$

or otherwise firm $j$ is better off breaking negotiations with its current manager and seeking a replacement in the labor market. Under a wide range of parameters values $w^\text{max}_{ij,t} \geq w^\text{min}_{ij,t}$ so that it’s possible for firm-manager pairs to strike a deal.

The wage of the manager from pair $(i, j)$ at date $t$ is therefore

$$w_{ij,t} = \eta_{ij} w_{ij,t}^\text{max} + (1 - \eta_{ij}) w_{ij,t}^\text{min}$$

---

$^2$ The risk-free equivalent outside wage delivers the same utility to the manager as the uncertain wage that she would get if she sought a new employer.
II.2.A. Equilibrium dynamics of industry wages under rational expectations

With rational participants who understand the wage setting process at individual firms, the estimates of the expected industry wage held by manager i and firm j of pair (i,j) are, respectively, the fixed points of the equations

\[
\mu_{i,t}^M = E_i(w_{i,t}) = \bar{\mu}_{i,t}^M + \left(1 - \bar{\mu}_{i,t}^M\right) E_i(w_{i,t}^{\text{min}}) = \mu_{i,t}^M + \lambda^F \left(1 - \lambda^M\right) \mu_{i,t}^M - \frac{1}{2} \left(1 - \lambda^M\right)^2 \sigma_{t,\text{predictive}}^2
\]  (5)

\[
\mu_{j,t}^M = E_j(w_{j,t}) = \bar{\mu}_{j,t}^M + \left(1 - \bar{\mu}_{j,t}^M\right) E_j(w_{j,t}^{\text{min}}) = \mu_{j,t}^M + \lambda^F \left(1 - \lambda^M\right) \mu_{j,t}^M - \frac{1}{2} \left(1 - \lambda^M\right)^2 \sigma_{t,\text{predictive}}^2
\]  (6)

Since, within every pair, the manager and the firm hold the same estimate \(\bar{\mu}_{i,t}^M\), we may collapse (5) and (6) into a single equation that is valid for both elements of the pair

\[
\mu_{ij,t} = \bar{\mu}_{ij,t}^M + \lambda^F \left(1 - \bar{\mu}_{ij,t}^M\right) \mu_{ij,t} - \frac{1}{2} \left(1 - \lambda^M\right)^2 \sigma_{t,\text{predictive}}^2
\]  (7)

Solving for \(\mu_{ij,t}\) yields:

\[
\mu_{ij,t} = \frac{\bar{\mu}_{ij,t}^M + \lambda^F \left(1 - \bar{\mu}_{ij,t}^M\right) \mu_{ij,t} - \frac{1}{2} \left(1 - \lambda^M\right)^2 \sigma_{t,\text{predictive}}^2}{\lambda^F}.
\]  (8)

We now can see that the heterogeneity in \(\mu_{ij,t}\), i.e., the heterogeneity in the estimates of the mean industry wage held by pairs, stems from heterogeneity in \(\bar{\mu}_{ij,t}^M\), i.e., from heterogeneity in the mean of the prior distribution of \(\mu_i\) held by pairs. Comparative statics of (8) show that \(\mu_{ij,t}\) is a positive function of \(\bar{\mu}_{ij,t}^M\) and \(\Pi\) and a negative function of \(\sigma_{t,\text{predictive}}^2\).

The variance of the predictive distribution of industry wages, \(\sigma_{t,\text{predictive}}^2\), is driven by two sources of uncertainty: first, the uncertainty stemming from the heterogeneity in the negotiation parameter \(\eta_{ij}\) in the population of possible firm-manager pairs, \(\sigma_{\eta}^2\); and second, the uncertainty about the population mean of \(\eta_{ij}\), \(\sigma_{\mu_{\eta,t}}^2\), that is perceived by firm-manager pairs. Since none of the two drivers is pair-specific, the predictive variance is the same across all pairs; moreover, only the second driver of uncertainty is time-indexed: hence, the time path of \(\sigma_{t,\text{predictive}}^2\) is determined by the time path of \(\sigma_{\mu_{\eta,t}}^2\).
We may now map the dynamics of industry wages once firms start to disclose publicly the wages paid to their managers.

Assume that firms disclose a noisy signal of the wage paid to their managers, i.e., they disclose

$$W_{ij,t} = w_{ij,t} + e_{ij,t} \quad (9)$$

where $e_{ij,t}$ is a pair-specific, iid, zero-mean random shock with variance $\sigma^2_e$, that is uncorrelated with all other variables of the model. As wage data becomes available, pairs update their estimates $\hat{\mu}^i_{\eta t}$ and $\sigma^2_{\eta t}$. With respect to the dynamics of $\hat{\mu}^i_{\eta t}$, we should observe a gradual convergence of the disparate initial estimates held by pairs towards the true mean of the distribution, $\mu_{\eta t}$. Simultaneously, the observation of wage data allows participants to improve the quality of their estimates of $\mu_{\eta t}$, leading to a reduction in $\sigma^2_{\eta t}$; with $\sigma^2_{\eta t}$ going down, the predictive variance, $\sigma^2_{\eta t,\text{predictive}}$, goes down too, raising the risk-free equivalent outside option of managers and, thereby, lifting the average industry wage.

In sum, the observation of wage data resulting from the mandatory disclosure of executive compensation reduces the variance while it increases the mean of industry wages. The latter effect results from the gradual shrinkage in the variance of the predictive distribution of industry wages held by participants, that occurs as they learn about parameter $\mu_{\eta t}$ from the observation of successive batches of disclosed wages.

The dynamics of equilibrium thus generate a ratchet effect in mean wage of industry executives following the mandated disclosure of executive compensation. Such ratchet effect should be felt strongly in the initial periods but taper off gradually over time. As more wage data becomes available, the reduction of $\sigma^2_{\eta t}$ brought about by the observation of additional batches of disclosed wages declines.

II.2.B. Equilibrium dynamics of industry wages under adaptive expectations

We now explore what happens if participants are either less rational or less well informed than in the benchmark case of rational expectations. Note that such benchmark requires participants to know the wage determination process at each individual firm and also to understand the implications of such process for the distribution of industry wages. We consider an alternative setup in which participants forecast the mean and the variance of industry wages by applying fixed updating rules to disclosed wage data. Our purpose here is to show that deviations from rational expectations can significantly extend the duration of the ratchet effect in the mean industry wage.

Suppose that participants develop estimates of the mean and variance of the predictive distribution of industry wages according to the updating rules

$$\mu_{ij,t} = \phi_m \bar{W}_{t-1}(1 - \phi_m)\mu_{ij,t-1} \quad (10)$$
\[ \sigma^2_{L,predictive} = \phi_v [\text{Var}(W_t) - \sigma^2_{e}] + (1 - \phi_v) \sigma^2_{0,predictive} \]  

(11)

where the initial estimates, \( \mu_{i,t} \) and \( \sigma^2_{0,predictive} \), are the same as those under rational expectations, to ensure that there participants have no forecast biases prior to the disclosure of wage data.\(^3\)

In this setup, participants forecast the mean and the variance of future industry wages as weighted averages of, respectively, past averages and past variances of disclosed wages, with the weights declining geometrically with the age of the sample (parameters \( \phi_m \) and \( \phi_v \) determine the rate of decay).

Let the first batch of wages disclosed after the mandated disclosure regime comes into force take place at \( t=0 \). Given our assumption of initial rational forecasts, we should observe \( \text{Var}(W_0) - \sigma^2_{e} = \sigma^2_{0,predictive} \). The observation of this first batch of disclosed wages initiates a process of convergence of the disparate estimates of the mean industry wage held by pairs, \( \mu_{i,t} \), leading to a gradual reduction in the variance of disclosed wages; hence, \( \text{Var}(W_t) < \text{Var}(W_0) \) and so the predictive variance held by participants at \( t=1 \) overestimates the realized variance, i.e.,

\[ \sigma^2_{L,predictive} = \phi_v [\text{Var}(W_0) - \sigma^2_{e}] + (1 - \phi_v) \sigma^2_{0,predictive} = \sigma^2_{0,predictive} > \text{Var}(W_1) - \sigma^2_{e} \]

This process goes on every period, with participants repeatedly overestimating the variance of disclosed wages. Because the predictive variance affects negatively the mean wage in the industry, such systematic overestimation of the predictive variance slows down the process of adjustment of the mean industry wage compared to the benchmark case of rational expectations.

But that’s not all. Once wages start to be disclosed, participants also underestimate the mean wage in the industry simply because wages are trending upward and thus estimates based on a weighted averages of past disclosed wages necessarily undershoot the target. Since realized wages depend positively on forecasted wages, the negative forecast bias in the mean wage entails lower wages than what would be the case if participants formed correct forecasts. In sum, adaptive expectations of the sort described in (10) and (11) generate systematic errors by participants in their forecasts of the mean and the variance of realized industry wages. Specifically, participants will underestimate the former and overestimate the latter.

The effect of forecast errors is to slow down the ratchet effect in mean wages. In other words, the mean wage will not rise so quickly as in the case of rational expectation but, in contrast, it will rise for much longer. Hence, a long lasting period of increasing wages could take place following the adoption of a mandated disclosure regime. The duration of this adjustment period of rising industry wages is negatively affected by the value of parameters \( \phi_m \) and \( \phi_v \). When these

\(^3\) The estimation equations (9) and (10) are appropriate if participants believe that the two first moments of the distribution of industry wages follow random walks. Different beliefs yield different estimation equations.
parameters are small, participants rely to a greater extent on stale samples of disclosed wages to build their forecast of future wages, thus making bigger forecast errors.

II.3.A. Simulations

To get a feeling for the duration of the ratchet effect under rational and under adaptive expectations we simulate the model. To be able to solve the rational expectations model analytically, we consider the special case in which $\lambda^M = \lambda^F = \lambda$. Under that specification, sample estimates of the unknown structural parameter $\mu_\eta$, can be directly obtained from disclosed wages.

With $\lambda^M = \lambda^F = \lambda$ equation (8) reduces to

$$\mu_{t,t} = \mu_\eta (1 - \lambda) \left(1 - \frac{1}{2} \right) \left(1 - \mu_\eta \right) \frac{(1-\lambda)^2}{\lambda} \sigma_{t,\text{predictive}}^2$$

(12)

The wage striken by pair $(i,j)$ at date $t$ is equal to

$$w_{i,j,t} = \mu_{i,j,t} \left(1 - \lambda\right) + \eta_{i,j} \lambda \Pi - \left(\frac{1}{2} \right) \left(1 - \eta_{i,j} \right) (1 - \lambda)^2 \sigma_{t,\text{predictive}}^2$$

(13)

Substituting (12) into (13) and simplifying yields

$$w_{i,j,t} = \left[\mu_\eta (1 - \lambda) + \eta_{i,j} \right] \left[\lambda \Pi + \left(\frac{1}{2} \right) (1 - \lambda)^2 \sigma_{t,\text{predictive}}^2\right] - \left(\frac{1}{2} \right) \left(1 - \lambda\right)^2 \sigma_{t,\text{predictive}}^2$$

(14)

so the predictive variance solves the second-order equation

$$\sigma_{t,\text{predictive}}^2 = \left[\lambda \Pi + \left(\frac{1}{2} \right) (1 - \lambda)^2 \sigma_{t,\text{predictive}}^2\right] \left[\left(\frac{1}{\lambda} \right) \sigma_{\mu_{t,t}}^2 + \sigma_\eta^2\right]$$

(15)

corresponding to the root that is increasing in the term $\left[\left(\frac{1}{\lambda} \right) \sigma_{\mu_{t,t}}^2 + \sigma_\eta^2\right]$.

On the other hand, if pairs, on average, hold unbiased expectations of market wages, the average disclosed wage is equal to

$$\bar{w}_t = \frac{\sum_{i=1}^{N} \hat{\mu}_{i,t} \Pi}{N} - \left(\frac{1}{2} \right) \left(1 - \frac{\sum_{i=1}^{N} \hat{\mu}_{i,t} \Pi}{N}\right) \left(1-\lambda\right)^2 \sigma_{t,\text{predictive}}^2 + \tilde{e}_t$$

(16)

where $N$ is the number of firm-manager pairs in the sample of disclosed wages.

Since $E \left(\sum_{i=1}^{N} \hat{\mu}_{i,t} \right) = \mu_\eta$, the average disclosed wage is a linear function of the unknown structural parameter $\mu_\eta$, plus a noise term. Participants may thus infer the unknown parameter $\mu_\eta$ from the
observation of the mean disclosed wage, \( \bar{W}_t \), simply by inverting the linear function. The estimate of \( \mu_\eta \) based on the sample of wages disclosed at date \( t \), taking into account that \( E(\varepsilon_t) = 0 \), is equal to

\[
\hat{\mu}_\eta^{\text{sample } t} = \frac{\bar{W}_t \left( \frac{1 - l^2}{\delta} \right) \sigma_{\text{predictive}}^2}{\frac{1}{N} \left( \frac{1 - l^2}{\delta} \right) \sigma_{\text{predictive}}^2}
\]

which has a variance equal to

\[
\text{Var}(\hat{\mu}_\eta^{\text{sample } t}) = \frac{\sigma_{\mu_\eta t}^2}{N} + \left[ \frac{1}{\frac{1}{N} \left( \frac{1 - l^2}{\delta} \right) \sigma_{\text{predictive}}^2} \right]^2 \frac{\sigma_{\varepsilon}^2}{N}
\]

Assume that the distribution of the noise in the wage disclosure process, \( e_{ij,t} \), as well as the prior distributions of the unknown parameter \( \mu_\eta \) held by participants are normal; then, applying standard Bayesian updating to conjugate normals, we obtain a posterior distribution of parameter \( \mu_\eta \) that is normal with mean and variance equal to, respectively,

\[
\hat{\mu}_\eta^{(i,t+1)} = \hat{\mu}_\eta^{(i,t)} \left[ \frac{\text{Var}(\hat{\mu}_\eta^{\text{sample } t})}{\text{Var}(\hat{\mu}_\eta^{\text{sample } t}) + \sigma_{\mu_\eta,t}^2} \right] + \hat{\mu}_\eta^{\text{sample } t} \left[ 1 - \frac{\text{Var}(\hat{\mu}_\eta^{\text{sample } t})}{\text{Var}(\hat{\mu}_\eta^{\text{sample } t}) + \sigma_{\mu_\eta,t}^2} \right]
\]

\[
\sigma_{\mu_\eta,t+1}^2 = \frac{\sigma_{\mu_\eta,t}^2 \text{Var}(\hat{\mu}_\eta^{\text{sample } t})}{\sigma_{\mu_\eta,t}^2 + \text{Var}(\hat{\mu}_\eta^{\text{sample } t})}
\]

We now have all the equations needed to simulate the model. We assume the parameter values contained in Table 1:

<table>
<thead>
<tr>
<th>N</th>
<th>( \mu_\eta )</th>
<th>( \sigma_{\mu_\eta}^2 )</th>
<th>( \sigma_{\varepsilon}^2 )</th>
<th>( \Lambda )</th>
<th>( \Pi )</th>
<th>( \sigma_{\mu_\eta,0}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0,5</td>
<td>0,1</td>
<td>1</td>
<td>0,4</td>
<td>3</td>
<td>0,2</td>
</tr>
</tbody>
</table>

Table 1. Simulation parameters

We simulate 10000 runs of the model, each for a duration of 10 periods following the adoption of the mandated disclosure regime at date \( t=0 \). Each run starts by drawing \( N \) values of \( \eta_{ij} \) from a normal distribution with mean \( \mu_\eta \) and variance \( \sigma_{\mu_\eta}^2 \) and \( N \) values of \( \mu_\eta^{(i,0)} \) from a normal distribution with mean \( \mu_\eta \) and variance \( \sigma_{\mu_\eta,0}^2 \). For the adaptive expectations model we consider three cases: (i) \( \phi_m = \phi_v = 1 \); (ii) \( \phi_m = \phi_v = 0,8 \); and (iii) \( \phi_m = \phi_v = 0,4 \).

Figure 1 and 2 report respectively the trajectories of the average industry wage and the variance of industry wages from the first batch of disclosed wages at \( t=0 \) to the 10th batch revealed at \( t=10 \).
Each point in the trajectories is computed as the average of the corresponding points generated by each of the 10000 runs.

Starting with rational expectations, we see that the average wage converges very quickly to the new equilibrium value: within two periods following mandated disclosure, over 95% of the upward adjustment has been completed. The clue to this quick adjustment process lies in Figure 2. With 20 firms disclosing the wage paid to their executives each period, the variance of the sample estimate of \( \mu_\eta \) is very small compared to the variance of the prior distribution of \( \mu_\eta \) held by participants. As a result, \( \sigma_{\mu_\eta,t}^2 \) converges very quickly to zero, pulling the predictive variance rapidly towards its new lower steady state value. In other words, the low variance of the sample estimates of \( \mu_\eta \) allows for a quick resolution of the uncertainty faced by participants about the exact value of the parameter once the mandated disclosure regime ramps up.

Under adaptive expectations, however, the adjustment process takes much longer. When \( \phi_m = \phi_v = 0.4, \), both the variance of wages and the average wage converge markedly slower than under rational expectations. When \( \phi_m = \phi_v = 0.8 \) or \( \phi_m = \phi_v = 1 \), the adjustment process in the variance of wages is undistinguishable from that under rational expectations, but the adjustment process in the average wage continues to be more sluggish. Even in the limit case where participants rely solely on the most recent batch of disclosed wages to build their forecasts of the mean and the variance of industry wages (i.e., the case \( \phi_m = \phi_v = 1 \)), the average wage takes much longer to converge to its new steady state than under rational expectations. We are led to conclude that under adaptive expectations the ratchet effect in the mean industry wage triggered by the disclosure of executive compensation will last for long.
Figure 1. Average industry wages following the mandated disclosure of executive pay. The trajectory of the average wage is simulated 10000 times; the figure reports, in each period subsequent to mandated disclosure, the average values across the 10000 trajectories.

Figure 2. Variance of industry wages following the mandated disclosure of executive pay. The trajectory of the variance of wages is simulated 10000 times; the figure reports, in each period subsequent to mandated disclosure, the average values across the 10000 trajectories.

IV. Conclusion

This paper develops a model that exploits the effect of mandated disclosure of executive compensation on the aggregate dynamics of CEO wages. It examines how transparency in CEO compensation practices of peer firms influences the determination of CEO pay at individual firms, exploiting the effect of the information revealed by transparency on the estimates held by executives and their shareholders about the value of their outside options when they bargain over CEO compensation. The key result is that when managers and shareholders learn about the properties of the distribution of CEO wages within their peer group from the observation of disclosed wages, the adoption of a mandated disclosure regime will trigger an escalation process in the mean compensation of CEOs and a simultaneous contraction process in the variance of compensation. The mean compensation rises while the variance declines until new batches of disclosed wages cease to provide new information about the wage distribution. The paper also
argues that the duration of the escalation in the mean wage depends on how managers and shareholders develop their forecasts of future wages: in a simulation we find that if managers and shareholders are rational and perfectly informed – i.e., they develop forecasts that are self-fulfilled given their beliefs and, additionally, fully understand the implications of the wage setting process of CEOs at individual firms on the aggregate distribution of wages – then the escalation process is quick and short. If, however, managers and shareholders develop forecast by updating previously held forecast to reflect the information revealed by newly disclosed wages using simple fixed rules, the process is slow and long.

The model leaves an important issue unresolved. Since it assumes away the roles played by managerial effort and moral hazard in the determination of optimal CEO compensation, focusing exclusively on the participations constraints of CEOs and shareholders, it is silent about the split between fixed and incentive pay (in the model there is only fixed pay). Edmands, Gabaix and Jenter (2017) document that the steady rise in total CEO pay that occurred in the last couple of decades was accompanied by an explosive growth in performance-related pay, in the form of stock option grants; from being a small portion of CEO pay in large-cap firms, stock option grants became the largest component of pay in the 1990s for such firms (after the stock market decline of 2000-01, stock options were increasingly replaced by restricted stock grants). As a result of these trends, the sensitivity of CEO wealth to stock price performance rose consistently between mid-1970s and the early 2000s, remaining high ever since. These authors argue that any theory that seeks to explain the surge in CEO total pay needs to account for the increasing importance of incentive compensation too. We leave that investigation for future research.
References


