Signaling within the firm

Ludovica Ciasullo

November 2024

Please click here for the latest version

Abstract

Promotions are a major driver of wage growth. I study how firms learn who to promote, when ability is hard to observe. Workers can exert costly effort in order to signal higher potential. Using Portuguese administrative data, I explore the role of a particular form of effort, overtime, in shaping careers in retail and hospitality. I develop a principal-agent model showing how overtime hours serve as a signaling device within firms. I test the model using a 2012 reform that quasi-exogenously reduced the overtime pay premium for some workers. Consistent with the model, overtime hours fall post-treatment and promoted workers are better selected. These findings are inconsistent with alternative leading theories of internal labor markets.

I am particularly indebted to Debraj Ray and Martin Rotemberg for invaluable guidance and advice throughout this project. I am especially grateful to Petra Moser and Jacob French, for their dedicated support and for their essential comments. I also thank Sandra Black, Jaroslav Borovička, Arielle Bernhardt, Nicolò Ceneri, Angela Crema, Tommaso De Santo, Raquel Fernández, Takao Kato, Anne Sophie Lassen, Robin Lenoir, Moritz Lubczyk, Erik Madsen, Sahar Parsa, Helena Pedrotti, Minu Philip, Kirtivardhan Singh, Teresa Steininger, Sharon Traiberman, Martina Uccioli, Michele Valinoti, Anna Vitali, Daniel Waldinger, Michael Waldman, and all the participants to NYU Applied Micro Lunch, and the NYU Gender Economics Workshop, for precious comments and suggestions.

Department of Economics, New York University. Email: lc4177@nyu.edu

1 Introduction

Promotions within the firm constitute both an important source of wage growth for workers (Huitfeldt et al., 2023) and a crucial strategic choice for firms (Lazear and Oyer, 2004). A large body of theoretical literature has studied such internal labor markets as an assignment problem in which employers learn whom to promote (Gibbons and Waldman, 1999): workers are typically heterogeneous in their ability and the firm acquires information on their unobserved characteristics before sorting workers into the most appropriate job (Rosen, 1985; Pastorino, 2024). The idea that workers may send signals of their ability to their employers to influence such learning process is very intuitive: if exerting effort in the workplace is costly, then the employer will think highly of those who do it. Yet, we don't have empirical evidence about whether signaling in the workplace happens in practice.[1] My paper is the first provide direct quasi-experimental evidence on the use of worker signaling.

Empirical tests of signaling in this context are hard to provide because a correlation between effort and future promotion can be rationalized with different models of the internal labor market. I exploit the idea that a variation in the direct monetary compensation for effort represents a variation in the signaling content of effort itself. I argue that the set of predictions that I theoretically derive (and empirically test) from a signaling model are not implied by other theories of internal labor markets.

I begin by developing a principal-agent model to investigate signaling within the firm. In particular, I extend the classical Spence (1973) signaling model by introducing a "transfer" (the overtime premium) that the receiver (the firm) pays to the sender (the worker) anytime the signal is chosen, and depending on the signal intensity. In my setting, worker heterogeneity is summarized by an unobserved parameter that represents the worker's ability if promoted. The firm's revenue depends on overtime hours in the first period and on the type of the promoted workers in the second period. The worker faces a cost for working overtime that depends on her type, but types are equally productive before promotion occurs. The worker chooses her overtime hours, and the firm decides whether to promote her. Because overtime hours, besides acting as a signal, are associated with a payment, different types of workers would want to choose different overtime hours, from a myopic perspective that ignores the promotion implications of their choice.

[1] Different settings with a similar structure, such as education, sick leaves, or maternity leave length, have been empirically investigated under the lenses of a signaling model (Lang and Kropp, 1986; Bedard, 2001; Heywood and Wei, 2004; Markussen, 2009; Clark and Martorell, 2014; Tô, 2018). Tô (2018) is particularly close since the paper evaluates whether firms use maternity leave length as a signal of mothers' disutility to work by developing a signaling model and empirically testing its predictions following a policy change that affects maximum leave length, and therefore ability to signal through leave.

I show that such myopic dispersion must translate, in the signaling model in which promotion concerns are introduced, into *some* degree of pooling in overtime hours across types in *every* equilibrium, which is in contrast to the classic signaling result of Spence (1973). In that case, absent the signaling model the workers would pool at the same, low-cost action, and it is the signaling game that induces a separation of types. In my case there is a discrete increase in payoff that the worker enjoys if she is promoted, but also a marginal return to each overtime hour given by the hourly wage: absent the signaling motive workers would pick different values of overtime hours, while they do instead pool once the promotion concern is considered.

By imposing a refinement in my continuous-action signaling game akin to the intuitive criterion, I show that just one equilibrium survives. Working with this unique equilibrium set of choices, which involves some pooling and some separation, I then examine the effect of a change in overtime pay. When that pay exogenously declines, I prove that the equilibrium changes so as to generate improved selection. This results creates scope for the empirical identification of a signaling component in the institution of promotions.

Specifically, the model generates testable predictions that link the signaling role of overtime to its piece-rate payment. In the unique equilibrium that I uncover, promotions are granted to workers who choose "enough" overtime hours, that is, above an endogenous equilibrium threshold that I can compute. Second, as the overtime pay premium goes down, this threshold decreases (fewer overtime hours are sufficient to obtain a promotion). Third, as the overtime pay premium goes down, the quality of the promoted workers increases.

The main intuition for the last result is the following. An increase in the cost of the signal (i.e., in overtime premium) has two different effects on the equilibrium allocation: on the one hand, the marginally promoted worker in the initial regime is now unwilling to supply the same number of overtime hours; on the other hand, this makes overtime hours more informative of the worker's ability and therefore strengthens the connection between overtime hours and promotions, making overtime hours more valuable for the worker. I argue that the equilibrium conditions, and in particular the fact that the firm is indifferent about promoting workers who pool at the threshold, imply that the marginally promoted type must increase.

Finally, I argue that my model's theoretical predictions are at odds with those of leading alternative theories of internal labor markets. First, if the relationship between overtime hours and promotions is about learning-by-doing, then one would not expect better promoted workers after overtime hours are reduced. Second, if such relationship is instead explained by incentive provision, then a decrease in the overtime wage would lead the firm to strengthen the hours-promotion correlation, but this would be at the expense of average quality of promoted workers.[2]

To document a role of signaling within the firm as a determinant of promotions, I test the predictions of my model on a sample of full-time workers in Portugal from 2006 to 2014.

I rely on administrative data that include, among other variables, the worker's position in an eight-rung hierarchy ladder. This ladder is defined nationally and is common to all firms, allowing me to unambiguously identify a promotion whenever a worker climbs one rung, from one year to the next. Crucially, the dataset reports overtime hours and overtime pay separately from their "regular" counterparts, hence allowing me to compute overtime hourly wage directly. The overtime premium is the percentage difference between this wage and the normal hourly wage.

I first show that overtime hours are correlated with future promotions. In particular, doing *any* overtime implies an increase in the probability of being promoted the following year of 9.2% relative to the baseline promotion probability for non-overtime workers of 9.5%. This effect is particularly concentrated among those who work one to four overtime hours a month, but does not increase linearly with overtime hours: consistently with the threshold structure of the equilibrium in my model, the shape of the relationship between overtime hours and future promotions is concave.

My identification is based on two key institutional details specific to Portugal. First, firms in Portugal have limited discretion in their overtime pay structure, as the employment conditions are largely determined through collective bargaining agreements. Second, A 2012 austerity reform brought down the mandated premium, defined as the percentage difference between the overtime wage and the normal wage, for *all* firms to a common value. This resulted in different decreases for different firms, based on their pre-reform levels of the overtime pay premium, which had been determined by collective bargaining, rather than by firms individually setting it optimally.

I show that fewer overtime hours are enough to be promoted after the reform: overtime hours decrease for all treated workers after the reform, and in particular for those who are about to be promoted, consistent with the model prediction of a lower overtime hours threshold.

The model predicts that selection through promotion improves after the reform, and indeed I document the positive impact of the reform on the quality of promoted workers.

[2] For a review of the theoretical and empirical literature that studies promotions as incentives, or motivational tools to encourage workers to exert effort, see Georgiadis (Forthcoming). The idea that promotions follow to human capital accumulation dates back to Becker (1962). See also, for example, Prendergast (1993) and Kwon and Milgrom (2014).

I compare workers whose latest promotion happened in the years 2009-2011 (before the reform) with those who are promoted at least once in the years 2012-2014 (after the reform), using a battery of proxies for worker quality and suitability for the firm. For example, as the pre-reform overtime wage goes from being equal to the normal hourly wage to being twice of it, the average annual wage growth of people who are promoted after the reform increases b 0.85 percentage points (26% of the baseline mean of 3.2%).

This result pins down the role of signaling within the firm, because it is inconsistent with two alternative leading theories of internal labor market, namely human capital accumulation and incentive provision. In the human capital accumulation view of promotions, workers are promoted when their ability has grown "enough": if overtime hours contribute to the worker's learning, then a decrease in overtime wage, and the consequent decrease in overtime hours, could not lead to increased quality of selection through promotions. In the incentive provision theory, promotions are a reward (alternative to the monetary direct payment) that incentivizes the worker to supply overtime hours. In this model, a decrease in overtime pay will lead to a strengthened relationship between overtime hours and promotion, because the firm will compensate the decreased monetary incentive with an increased promotion incentive to maintain the workers engaged in overtime work. Such more generous promise of promotion would lead to a decline in the average ability of promoted workers.

Overall, the empirical analysis confirms the model predictions and is in line with the firm operating at least in part with a selection motive, whereby overtime is used to identify the worker's type.

Related Literature I contribute to two strands of the organizational economics literature, by providing empirical evidence of the role of worker's signaling in determining careers within firms, and by documenting a causal relationship between long working hours and promotions.

A robust empirical and theoretical literature that studies asymmetric employer learning assumes that workers are "passive agents", unable to affect the firm's belief about their unobserved heterogeneity (Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lluis, 2005; Hunnes, 2012; Kahn and Lange, 2014; Wu, 2024; Pastorino, 2024).[3] I focus on entry-level jobs, and routine tasks, a context where current productivity differences are less likely to matter, and worker choice of effort is probably very relevant.[4]

[3]For example, in (Wu, 2024) employers learn about their worker heterogeneous talent by observing occurrence of exceptional achievements, but the worker cannot influence how likely it is that such breakthrough happen.

[4]Acemoglu et al. (2011) argue that "'the core job tasks of these occupations follow precise, well-

My paper is the first to provides both empirical evidence and theoretical justification in favor of worker's signaling in determining careers within the firm.[5] Closely related is Ghosh and Waldman (2010), who also model asymmetric learning in a context that allows for worker's choice of effort, but who focus on optimal promotion policy in a context where promotions are a signal, to other firms, of the worker's ability.

Importantly, the existing literature on employer learning is mostly silent on the possible role of incentive provision as an alternative explanation for the observed promotion patterns.[6] Empirically, distinguishing between a model where promotions provide the incentive to exert effort to a set of heterogeneous workers, and a model in which promotions follow the firm's realization of the actual ability of the worker, is challenging (Treble et al., 2001; Kahn and Lange, 2014) and I am unaware of papers that address this distinction.[7] I use my model to derive a set of predictions under the signaling scenarios and discuss extensively why incentive provision alone cannot generate the observed patterns.[8]

The second strand to which my paper contributes is one that studies the relationship between long working hours and promotion. My paper is, to the best of my knowledge, the first to provide quasi-experimental evidence of this relationship, and in particular of worker signaling as the underlying mechanism. The most closely related paper is Frederiksen et al. (2024): using detailed administrative data, matched to survey data, they use information on working hours (including in which firm they are performed, when in the day/week, and how they relate to peers' hours) to study how they impact the probability of promotions from management to top management positions. Their findings are

understood procedures", which speaks to the idea of output being directly related to measurable inputs rather than intangible individual differences.

[5]Within the employer learning literature, an important distinction concerns whether the information acquired by the firm are simultaneously observed by other potential employers (symmetric learning) or limited to the current employer (asymmetric learning). My paper belongs to the second strand as I assume that only the current employer can observe effort (as proxied by overtime hours). This is consistent with Bossler and Grunau (2020) who show that overtime hours are correlated with internal promotions but not with career moves across firms. There is robust empirical evidence that supports asymmetric learning: notable examples include Gibbons and Katz (1991), Pinkston (2009), Kahn (2013), and Bognanno and Melero (2016). On this, see also Waldman and Yin (2024).

[6]Furthermore, most papers that argue the importance of employers' learning in worker wage growth exploit data from single firms. Some exceptions use administrative data (Lima and Telhado Pereira, 2003; Van der Klaauw and Dias da Silva, 2011) and survey data (Lluis, 2005) to test this prediction, but they do not rely on quasi-experimental variation. Relying on one-firm data has the advantage of providing detailed information on workers' output, or performance evaluations, at the cost of compromising on external validity, as the firms studied tend to be very large, and generally not representative.

[7]One exception is Pastorino (2024) who briefly engages, in an Appendix, with the possibility that performance incentives through promotion tournaments may be an alternative explanation for the patterns she explains through employers learning.

[8]We also have evidence that employer learning alone cannot explain promotion: Benson et al. (2019) show that promotions don't always go to the best potential manager and that instead they are at least in part used to provide incentives in the current job.

consistent with accumulation of firm-specific human capital, as well as with signaling of ability through long working hours to the present employer. My work complements theirs by providing causal evidence in favor of a signaling mechanism. Other relevant papers in this literature include Bell and Freeman (2001); Meyer and Wallette (2005); Pannenberg (2005); Bratti and Staffolani (2007); Anger (2008). I add to this literature in three different ways. First, I rely on a policy experiment (and not only on cross-sectional variation) to argue that worker effort, measured through overtime hours, acts a signal for the employer. Second, I use administrative data (rather than survey evidence) in which overtime hours are directly recorded (hence reducing the potential measurement error implied by imputation). Third, I explicitly consider multiple alternative theories and the extent to which they could rationalize the observed empirical patterns.

The policy I study has been examined only by Martins (2017), who finds that firms what cut overtime premiums after the 2012 reform increased their use of overtime and experienced higher employment and sales. This result compares firms based on whether or not they reduced premiums, while my approach exploits variation in pre-reform overtime premiums, which strongly predicts the size of the drop and is arguably exogenous.

In terms of its theoretical contribution, my paper belongs to a literature focused on career concerns (Landers et al., 1996; Schöttner and Thiele, 2010; Barlevy and Neal, 2019). My model uses the structure of the classic signaling framework of Spence (1973) to study internal labor markets. I augment this framework with the assumption that the signal (overtime hours) is compensated with a transfer (overtime pay): this, together with the discrete increase in payoff upon promotion, implies that separation of types cannot occur in equilibrium. In the case of Spence (1973), the introduction of the signaling motive allows the principal to distinguish different types, while in my case subordinating promotions to an independently rewarded signal creates pooling that would otherwise not be observed. While the game I construct has infinitely many equilibrium, a relatively mild version of the intuitive criterion effectively disciplines it, allowing me to prove the existence of a unique equilibrium.

The remainder of the paper is organized as follows. Section 2 outlines the theoretical framework, it derives both the unique equilibrium and the comparative statics predictions about a change in overtime wage, and it discusses alternative theoretical understanding of internal labor markets. Section 3 describes the institutional context in which my empirical analysis is based, section 4 details the dataset I rely on and the empirical strategy I employ, section 5 presents my results, and section 6 concludes.

2 Model

2.1 Setting

A firm employs a worker for two periods, and it can promote her at the end of the first one. In period 1, all workers work a fixed number of normal hours,[9] and earn an exogenous hourly wage. Additionally, they can freely choose overtime hours at some exogenously given hourly wage. I normalize the workers utility from working normal hours, and the firm's profits from it, to 0. The marginal product of each overtime hour is 1. Overtime hours h are freely chosen by the worker and are paid at the rate $w < 1$. In short, this formulation imposes that overtime is always valuable to the firm, and that there is no cap on how many overtime hours a worker can work.

The worker is characterized by a type, denoted by θ, that represents her ability or productivity upon promotion. All worker types are otherwise identical, and their baseline output does not allow them to directly signal their future productivity differences. Ability θ, which is unobserved during the initial period, is distributed according to the probability density function g on the support $[\underline{\theta}, \bar{\theta}]$.

The information structure is the following. At the beginning of the period, workers observe their type θ and choose their overtime hours h. At the end of the first period the firm, to which the worker's type θ is unknown, observes the worker's choice of overtime hours and selects a promotion probability $\pi(h)$.

Worker's cost of working overtime is $c(h, \theta)$, which is increasing and convex in h, decreasing in θ, and with marginal cost also decreasing in θ.[10] Then, the firm, having observed h, decides whether to promote the worker or not. If they do, their profit will depend on the worker type, but they'll need to pay a higher wage, as specified below. If the firm does not promote the worker, its profits are equal to 0.

The underlying reason for the correlation between ability if promoted and reduced cost of overtime is not modeled. However, I claim that such an assumption is reasonable from at least two perspectives. First, it could be that as the worker advances in the firm hierarchy, her tasks change in the direction of requiring more hours: it is reasonable to assume that supervisors need to work more than blue-collar workers, and managers more than supervisors. If that's the case, then the selection of whom to promote requires, by definition, identifying who is willing and/or able to supply the quantity of labor required

[9]This is consistent with the data, as the distribution of normal hours (excluding overtime) shows a mass at the weekly limit of 40 hours a week

[10]Formally, $c'(h, \cdot) > 0$ (cost of working is increasing in hours worked), $c''(h, \cdot) > 0$ (cost of working is convex in hours), $c'(\cdot, \theta) < 0$ (cost of working is decreasing in θ) and $\frac{\partial}{\partial \theta} \frac{\partial}{\partial h} c(h, \theta) < 0$ (as θ increases, the cost of working is "less increasing" in hours).

for the upper-level job. However, even if we do not want to conceptualize the upper-level job as inherently requiring long hours, my assumption still holds if some underlying personality trait of the worker governs both their value to the firm as a supervisor or manager, and their ability to supply long overtime hours. Loyalty to the firm, for example, or commitment to one's career, are subjective traits that reasonably impact how valuable a promoted worker can be to the firm. At the same time, these characteristics arguably reduce the utility cost of working long hours.

A promoted worker is paid a surplus of V from the next period on, and contributes a *net* amount θ to the firm payoff.[11] The firm will therefore earn negative profits if they promote worker $\theta < 0$: I assume that the type $\theta = 0$ that makes the firm indifferent about its promotion decision is interior to the support of θ. Workers who are not promoted stay in the same job, and they earn and produce 0.

Suppose for a moment that promotion is not a concern. Then, given my assumptions, there is a unique myopic first best value of overtime hours for each θ, i.e. a choice of overtime hours h that maximizes period-1 worker payoff. Call this value $h^m(\theta)$: we take it to be strictly positive for every θ, so it is characterized by the first-order condition of the isolated period-1 problem, i.e.

$$\frac{\partial}{\partial h} c(h,\theta)\Big|_{h=h^m(\theta)} = w.$$

In addition to this myopic consideration, there are the prospects of promotion. From the worker's perspective, this is described by a probability $\pi(h)$ for every observed choice of h.

Therefore, her overall payoff at the beginning of stage 1, when she is choosing overtime hours, is given by

$$\underbrace{wh - c(h,\theta)}_{\text{period 1}} + \underbrace{V\pi(h) + (1-\pi(h))0}_{\text{period 2}}$$

We reiterate that the worker freely chooses h, in line with our assumption that overtime is always valuable to the firm.

Conditional on observing overtime hours h and promoting the worker, the firm's expected period 2 profit is simply given by $\mathbb{E}(\theta \mid h)$. Bayes' rule will be used whenever possible to compute this expectation. It is immediate that the firm will promote the worker for sure if his conditional expectation is positive, never promote if it is negative, and will be indifferent (and therefore can promote with any chosen probability $\pi(h)$) if this conditional expectation is precisely zero.

[11]That is, a promoted worker can be viewed as generating a gross output of $V + \theta$.

There are some basic differences between my model and the classic signaling game from Spence (1973). Because overtime is compensated, the optimal hours will vary with the type of worker even when they are chosen without heed to promotion. This, together with the fact that the worker's payoff increases discontinuously upon promotion, implies that some degree of *pooling* is necessary in equilibrium, contrary to the case of Spence (1973) where separation is achieved via the signaling process. [12]

I also carry out a comparative statics exercise that has no immediate counterpart in the standard signaling model, i.e. changing the overtime rate.

2.2 Equilibrium

In this section, I define an equilibrium of this game, explain why there are multiple equilibria and employ a refinement that guarantees uniqueness.

An equilibrium of this signaling game consists of three objects:

1. a map from any possible vale of h to the firm's posterior belief about θ, computed using Bayes' rule whenever possible;

2. a promotion decision $\pi(h) \in [0,1]$, which maximizes the firm's expected product in the second period given the posterior belief on θ derived from observing h;

3. a choice of overtime hours $h(\theta)$ for each type, which is a best-response to the firm's promotion decision;

There are multiple equilibria to this game, indexed by different sets of off-path beliefs (expectations over θ that would be generated by overtime hours that are not chosen in equilibrium). For instance, suppose that an equilibrium has two pooling sets: all types $\theta \in [\theta_1, \theta_2]$ choose h_1, and are promoted with some positive probability $\pi(h_1) < 1$, and all types $\theta \in [\theta_3, \theta_4]$, with $\theta_3 > \theta_2$ choose $h_2 > h_1$, and are promoted with certainty. For this to be an equilibrium, as Appendix A explains in detail, it must be that by choosing some $h' \in (h_1, h_2)$, type θ_3 loses the promotion advantage, which requires that $\mathbb{E}(\theta \mid h') \leq 0$. Hence, it must be that, upon observing h_1 the firm's posterior belief over θ is weakly positive (because there is a positive promotion probability associated to this choice), but that upon observing $h' > h_1$ the firm's belief is that the worker type is weakly negative. So either $\mathbb{E}(\theta \mid h_1) > \mathbb{E}(\theta \mid h')$, or $\mathbb{E}(\theta \mid h_1) = \mathbb{E}(\theta \mid h') = 0$. In the former case, the firm believes that a negative type would choose higher overtime hours than a positive type, which is

[12]Another difference with the setting of Spence (1973) is that in my model the promotion payoff is independent of the worker type.

intuitively implausible, but possible as long as the value h' is not chosen in equilibrium. In the latter case, if $\mathbb{E}(\theta \mid h_1) = 0$, then it must be that $0 \in [\theta_1, \theta_2]$, and therefore by observing $h' > h_1$ the firm needs to think that the type who deviates is internal to the interval $[\theta_1, \theta_2]$: since such deviation does not lead to an increase in the promotion probability, this belief is implausible because the $h^m(0)$ must be lower than $h^m(\theta_2)$, so that if type 0 was deviating to h', then θ_2 should want to do the same. Infinitely many equilibria, all supported by different beliefs, belong to this category.

I impose a version of the intuitive criterion from Cho and Kreps (1987), as formulated in Esteban and Ray (2006), that refines the equilibrium concept.

Definition 1 (No-equilibrium-dominance) *For any given equilibrium, and any type θ say that a choice of overtime $h' \neq h(\theta)$ is equilibrium dominated for θ if*

$$wh' - c(h', \theta) + V < wh(\theta) - c(h(\theta), \theta) + \pi(h(\theta))V.$$

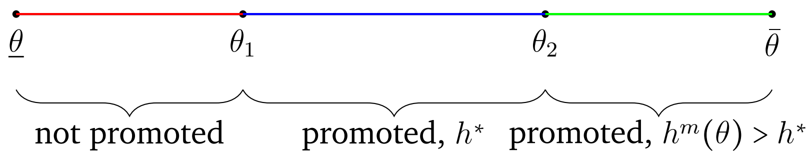
In words, h' is equilibrium dominated for worker θ if the equilibrium payoff associated to $h(\theta)$ is higher than what she would get by working h' overtime hours and being promoted with certainty. Under no-equilibrium-dominance, the firm, after observing an off-path choice of overtime hours, h', forms a posterior whose support excludes any type for which h' is equilibrium dominated.

The following proposition characterizes the unique equilibrium of the refined version of this game.

Proposition 1 *Under no-equilibrium-dominance, there exists a unique equilibrium, characterized by a threshold for overtime hours h^*, and a pooling set of types $[\theta_1, \theta_2]$ such that:*

1. $\pi(h) = 1$ *if and only if* $h \geq h^*$*;*
2. $\forall \theta \in [\theta_1, \theta_2]$*,* $h(\theta) = h^*$*;*
3. $\forall \theta \notin [\theta_1, \theta_2]$*,* $h(\theta) = h^m(\theta)$*;*
4. $h^m(\theta_2) = h^*$*;*
5. $\mathbb{E}(\theta \mid h^*) = 0$

In equilibrium, all workers below θ_1 choose their myopic optimum, which is below h^*, and are therefore not promoted. All workers from θ_1 to θ_2 jump up to the choice h^*, which is above their myopic optimum, except for θ_2 for whom it exactly equals the myopic optimum. [13] All workers above θ_2 choose $h^m(\theta) > h^*$, i.e. they would have chosen the same overtime hours even absent the promotion motive.



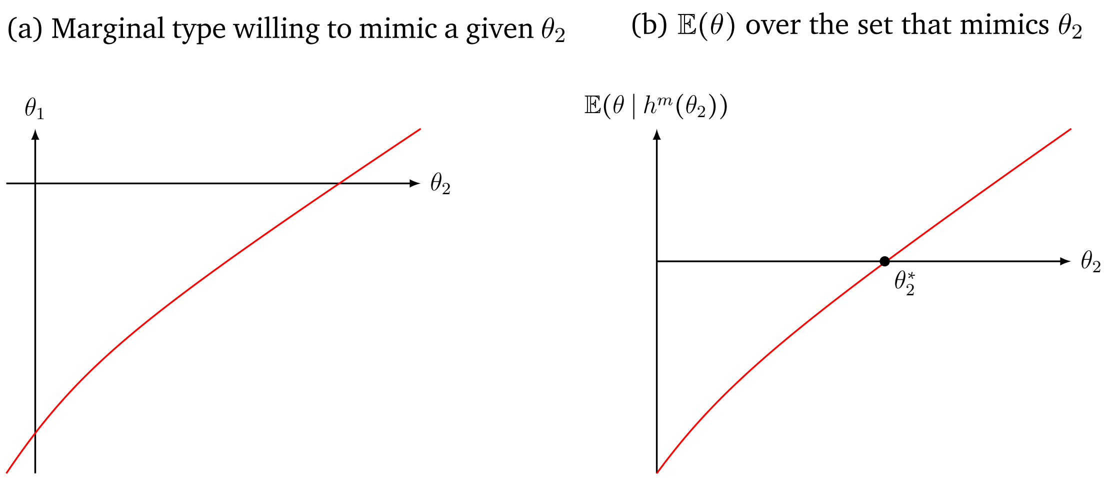
The proof of the proposition is in Appendix A. First, the proof establishes that in any equilibrium of the unrefined game there is at least one pool, i.e. a set of types who all chose the same overtime hours. The main intuition is that, absent any pool, all workers would choose different actions and therefore the firm's posterior belief about θ, for all the equilibrium choices of overtime hours, will be degenerate: since the firm wants to promote workers whenever $\mathbb{E}(\theta \mid h) > 0$, a marginally non-promoted worker would get a jump in the promotion probability from 0 to 1 by increasing their overtime hours, and the proof argues that this must be a profitable deviation for θ negative, but close enough to 0. Next, I show that in any refined equilibrium there is at most one pool. Imagine an equilibrium with two pooling sets, with promotion probability strictly positive on both, and increasing across the two sets. The intuition here relies on identifying a profitable deviation for a type θ that is interior to the second pool, the one that does not include 0: agent θ is pooling with other agents and not selecting her myopic optimum, but a small deviation towards $h^m(\theta)$ will be equilibrium dominated for type 0 and therefore, by the equilibrium refinement, type θ's promotion probability will still be equal to 1 under this deviation, hence contradicting the possibility of two pooling intervals. Finally, the proof characterizes the one-pool equilibrium, by pinning down the interval $[\theta_1, \theta_2]$ using the following conditions: $\mathbb{E}(\theta \mid h^*) = 0$ and θ_2 is the agent whose myopic first best coincides with the pooling action h^*.

The two diagrams below represent the two main forces that define the equilibrium. Panel (a) represents, for each value of θ_2, the corresponding value θ_1, uniquely pinned down by the indifference condition:

$$wh^m(\theta_1) - c(h^m(\theta_1, \theta_1) = wh^m(\theta_2) - c(h^m(\theta_2, \theta_1) + V.$$

[13] This formulation of Proposition 1 relies on assuming that the support of θ is "large enough". In particular, if $\theta_2 = \bar{\theta}$, then $h^m(\theta_2) < h^*$. In other words, I am assuming that $\mathbb{E}(\theta \mid h^m(\theta_)$ is equal to 0 for a value of θ_2 that is internal to the support of θ. The rest of the Proposition is unchanged. See Appendix C for more details on the boundary case.

The implicit function $\theta_1(\theta_2)$ will be increasing: a higher θ_2 will have a higher myopic first best, so that the type that is marginally willing to mimic it will have to be higher. Panel (b) represents how a specific θ_2 is pinned down by the firm's indifference condition: the expected value of θ over the pooling set needs to be equal to 0: as h^* increases, so does θ_1 (a higher action is harder to mimic), which means that the interval over which the expectation is taken shifts to the right, so that the expected value of θ over the pooling interval must increase. If $h^* = h^m(0)$, it must be that $\theta_1 < 0$, and we assume that the support of θ is large enough that if $h^* = h^m(\bar{\theta})$, then $\mathbb{E}(\theta \mid h^*) > 0$, which implies the existence of a unique value θ^* such that $\mathbb{E}(\theta \mid h^m(\theta^*) = 0$.[14]



2.3 A Decrease in the Overtime Premium

In this Section, I will outline the main predictions that the model delivers when carrying out a comparative statics exercise focused on a decrease in the overtime pay premium.

Proposition 2 *As overtime wage w decreases, the semi-pooling equilibrium results in lower overtime threshold for promotion h^*. Furthermore, if the convexity of the cost function increases in h and does not decrease in θ, then average quality of promoted workers increases. Formally, as w decreases, $\mathbb{E}(\theta \mid \text{promoted}) = \mathbb{E}(\theta \mid h \geq h^*) = \mathbb{E}(\theta \mid \theta \geq \theta_1)$ increases, provided that $\frac{\partial}{\partial^2 h} c(h, \theta) < 1$.*

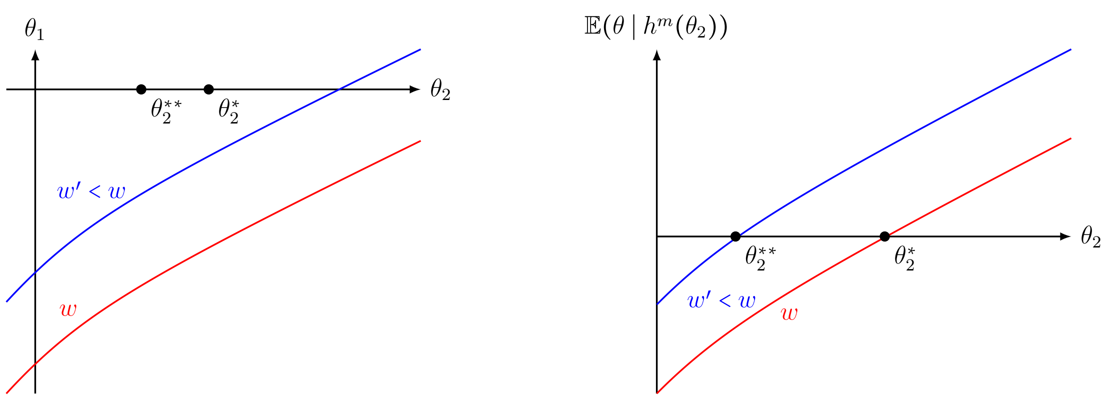
The proof is detailed in Appendix B, but the intuition is represented in the diagram below. Panel (a) depicts the relationship between θ_1 and θ_2, where the former is defined

[14]This assumption does not affect the core of my argument. See Appendix C for a discussion of the case where it doesn't hold.

as the lowest type who is willing to mimic the latter: as w decreases, and each overtime hour is paid less, the marginal type who is willing to mimic a fixed $h^m(\theta_2)$ shift up, but at the same time a given θ_2 will have a lower $h^m(\theta_2)$ as her myopic optimum. In principle, it is ambiguous whether the first, direct effect of the change in overtime wage on the quality of the selection (any given h^* is harder to mimic) dominates the second, indirect one (the h^* required to mimic is lower). What determines the net effect is the shape of the cost function, and in particular whether the high type θ_2 in the initial equilibrium is less responsive than the low type to the change in w, so that the change in h^* is not "too large". If this is the case, then the relationship between θ_1 and θ_2 in the new regime is represented by the blue line in panel (a) in the diagram below: for a given θ_2, as w decreases the lowest type who is willing to mimic θ_2 in exchange for promotion is higher. A sufficient set of conditions that generate this implication is that $c'''(h, \cdot) > 0$ and $\frac{\partial}{\partial \theta} c''(h, \cdot) \geq 0$. As an example of a class of cost functions that meets these conditions, together with the standard assumptions described in section 2.1, consider $c(h, \theta) = f(h) + g(\theta)h$, for $f(h)$ positive, increasing, convex and increasing convex, and $g(\theta)$ positive and decreasing.

Panel (b) below depicts the relationship between $\mathbb{E}(\theta \mid h^m(\theta_2))$ and θ_2: as w decreases, for a fixed θ_2 the marginal type that is willing to mimic it will be higher (as discussed above, and as depicted in the panel (b)), so the lower extreme of the interval over which the expectation $\mathbb{E}(\theta \mid h^m(\theta_2))$ is taken will be shifted up,for a fixed θ_2 (blue line in panel (b)). This implies that the value of $h^m(\theta_2)$ for which the expectation is 0 is lower, and therefore that the pooling action is lower. Since θ_2 is defined as the type such that $h^m(\theta_2) = h^*$, this means that θ_2 is lower. By the equilibrium characterization provided in Proposition 1, we need $\mathbb{E}(\theta \mid h^*) = \mathbb{E}(\theta \mid \theta \in [\theta_1, \theta_2]) = 0$. Since θ_2 has gone down, and the expected value over the interval needs to be equal to zero in the new equilibrium as well, that θ_1 needs to increase.

(a) Marginal type willing to mimic a given θ_2 (b) $\mathbb{E}(\theta)$ over the set that mimics θ_2



To recap my argument, I first establish that as w goes down, the lowest type who is willing to mimic θ_2^*, for a fixed θ_2^*, is higher (despite the fact that the myopic optimum of θ_2^* has decreased). Because of this, $\mathbb{E}(\theta \mid h^m(\theta_2^*)$ increases as w increases (the lower bound of the interval over which the expectation is taken has shifted up). Since $\mathbb{E}(\theta \mid h^m(\theta_2)$ is an increasing function of θ_2, and in equilibrium such expectation needs to be zero, it means that θ_2 decreases to θ_2^{**}. This will impact where type θ_1 sits, but θ_2^{**} is defined so that $\mathbb{E}(\theta \mid h^m(\theta_2^{**}) = 0$, which implies that θ_1 must increase, regardless of the shape of the distribution of θ: since $0 \in [\theta_1, \theta_2]$ for any equilibrium, and since in the new equilibrium θ_2 decreases, θ_1 must increase to preserve the constant expectation equal to 0.

2.4 Discussion

There are three testable implications that follow from my model.

Prediction 1 *Overtime hours are correlated with future promotions. In particular, for any set of firm-specific unobservable thresholds h^* the data generated by my model would show a concave relationship between overtime hours and future promotion probability.*

The first part of Prediction 1 follows directly from Proposition 1: since promotions are granted any time overtime hours are above a threshold h^*, and only in those cases, overtime hours are positively correlated with future promotion. For a given firm, increasing overtime hours beyond the threshold won't generate any further increases in the promotion probability, since in the equilibrium I derived in Proposition 1 the promotion probability is equal to 1 for any value of $h \geq h^*$. The second part of Prediction 1 follows from thinking of a cross-sectional datasets: different firms are in different equilibria, each characterized by the same threshold structure. On average, we observe a positive correlation between overtime hours and future promotions that displays a concave structure: at first, increasing overtime hours generates increases in the probability of future promotions, since low levels of overtime hours are below the unobserved h^*'s; higher values of overtime hours will be above more and more thresholds h^*, until the point where increasing them further does not boost promotion probability in any of the firms in the data.

Prediction 2 *As the overtime pay premium goes down, the promotion threshold in overtime hours decreases, and fewer overtime hours are sufficient to obtain a promotion.*

This prediction follows directly from Proposition 2, which argues that as w decreases, so does h^*. In the new equilibrium, for the lower overtime premium, all the workers whose equilibrium choice of overtime hours coincide with their myopic first best will do fewer overtime hours. This is because the myopic first best is defined as the choice at which the marginal cost of an additional hour equals the overtime wage w. Beside this effect, the workers in the pooling set also decrease their overtime hours, because the threshold they have to reach is lower. Empirically, I'll show overtime hours decrease, and in particular they decrease for people who are about to be promoted.

Prediction 3 *As the overtime pay premium goes down, the average type of the promoted people improves.*

This prediction is implied Proposition 2, which states that the marginally promoted type θ_1 is higher when the overtime premium is lower. Since the unobserved type θ represents, given my assumption, the worker's suitability for a good career, the data generated by this model would display better careers for workers who are promoted after the reform, compared to workers who are promoted before the reform.

3 Institutional Setting: Overtime Regulation in Portugal

This Section describes the institutional setting in which I study the signaling role of overtime. First I illustrate the overtime legislation before 2012, and then I explain the content of the reform which acts as quasi-experimental variation in my analysis.

The Status Quo Before 2012 Overtime work in Portugal is regulated by a series of articles in the Labor Code, and in particular its compensation is established by Article 268.[15] This article states that, absent any collective bargaining agreement that determines differently, overtime work is paid at the hourly rate with the following additions:

1. 50% for the first hour or fraction thereof and 75% per subsequent hour or fraction thereof, on a working day;

2. 100% for each hour or fraction thereof, on a mandatory or complementary weekly rest day, or on a public holiday.

The same article entitles collective labor regulation instruments to alter these premiums.

[15] The text of the Labor Code as introduced in 2009, Lei n. 7/2009, can be found at this link: `https://diariodarepublica.pt/dr/legislacao-consolidada/lei/2009-34546475`. The text of the previous version of the Labor Code, Lei n. 99/2003, can be found here: `https://diariodarepublica.pt/dr/detalhe/lei/99-2003-632906`. Article 258 of the 2003 version of the Labor Code had virtually the same text.

In practice, the law establishes default rates for the overtime premium, and defers to collective bargaining the possibility of modifying such rates. None of the registered contracts actually decreased the premium, so that the premiums stated in the Labor Code act as minimum levels (Martins, 2017).

The 2012 Reform The law that modifies the overtime legislation in Portugal has been introduced in accordance to the commitments of a Memorandum signed in May 2011 by the Portuguese government, under the rationale of "contain[ing] employment fluctuations over the cycle, better accommodat[ing] differences in work patterns across sectors and firms, and enhanc[ing] firms' competitiveness" (International Monetary Fund , 2011).[16]

The updated text of Article 268, as modified by Law 23/2012, [17] halves all the three premiums indicated in the Labor Code, leaving the rest of the overtime legislation unaltered. Hence, starting from August 1st 2012,[18] overtime work is established to be paid 25% more than the normal hourly wage for the first hour after the end of the normal working day, 37.5% more for any subsequent hour, and 50% more for any hour performed on a rest day or a holiday. Importantly, article 7 of the same law establishes that "the provisions of collective bargaining instruments and clauses of employment contracts, which came into force before August 1, 2012, and which provide for [...] increases in payment for overtime higher than those established by the Labor Code [...] are suspended until December 31, 2014."

Hence, *de facto* Law 23/2012 establishes new, lower premiums, and enforces them for two years. The heterogeneity in pre-2012 overtime premium levels, versus a common new value, determines that the drop varies in size across different firms.

The structure of overtime payment is summarized below:

Overtime hours in a day	Premium		
	Before 2012	2012-2014	After 2014
0	0	0	0
1	≥ 50%	= 25%	≥ 25%
2 +	≥ 75%	= 37.55%	≥ 37.5v%
on holidays	≥ 100%	= 50%	≥ 50%

[16]See Carvalho and Ribeiro (2022) for a broad discussions of the provisions of the Memorandum. The Memorandum was part of an economic and financial adjustment program funded by the European Union, the European Central Bank and International Monetary Fund.

[17]The Law can be read in its entirety at this link: `https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=1755&tabela=leis&so_miolo=`.

[18]Law 23/2012 was promulgated in June 2012 and set to come into force "come into force on the first day of the second month following its publication".

Figure 3 shows how the economy's average overtime premium evolved from 2006 to 2021 in the full sample of retail and hospitality workers in Portugal:[19] there is a sizeable drop in 2012, and the average overtime premium stays approximately constant for the subsequent two years, before going back up to the pre-reform levels once collective bargaining was again allowed to alter it.[20] The fact that the overtime pay premium goes back to the pre-reform levels once the freeze on collective bargaining is lifted speaks to the exogeneity of the reform, which was altering an otherwise "absorbing" state.

4 Data and Empirical Strategy

In this section I describe the primary data source, an administrative matched employer-employee dataset. Next, I present the empirical strategy I employ to identify the effects of the reform.

4.1 Data

The matched employer-employee dataset I use is called *Quadros de Pessoal* (Personnel Records), it is collected by the Portuguese Ministry of Employment, and it includes information on all private sector firms that employ paid labor in the reference month of any given year.[21] Each year (in October), all Portuguese firms fill in a written survey, with information about the firm, the establishment, and the workforce.[22] Filling out the survey is mandatory, and the information the firm submits is required to be posted in a public space in the establishment, available for workers' consultation. The purpose of this data collection is independent from tax purposes, and instead it serves to verify firm's compliance with Labor Law and collective bargaining agreements (Cardoso and Portugal, 2005). The Autoridade para as Condições do Trabalho (ACT, Authority for Labor Conditions) is responsible for audits and inspections of firms.

The resulting dataset is a matched employer-employee panel that includes information on the firm's location, industry, sales volume, and age, as well as the worker demographics,

[19] Overtime premium is defined as the percentage difference in hourly wage between normal hours and overtime hours. For the construction of this variable in the data I use, refer to Section 4.1.2.

[20] From the time series plotted in Figure 3, we see that the average overtime pay premium increases from 90% to 100% around 2009-2010: this may be because when the Labor Code was introduced (in February of 2009), while the values of minimum overtime pay premium were unchanged, the text of the law was slightly modified to mandate the increase in hourly wage starting for the first hour *or fraction thereof* after the normal working day. Since may data only includes round numbers of overtime hours, I cannot verify whether this explains the observed small jump.

[21] Public sector and independent contractors are exempted.

[22] Worker-level information is provide for the entire workforce, including non-wage earners

occupation, tenure in the firm, collective bargaining contract, and hours and earning in the reference month.

The crucial feature of this dataset is that firms are required to separately record, for each worker, overtime hours and earnings separately for their "normal" counterparts. Monthly regular hours and associated payments are reported for all workers, and correspond to the actual hours worked (and corresponding earnings) that are not covered by the overtime regulation, i.e. within the regular working hours. For workers who do not work any overtime in October of a given year, the variables for overtime hours and overtime payments are be recorded as zero. Conversely, for all workers who work overtime, the firm reports the total number of hours and the total payments associated to those hours. This is a unique characteristic of this dataset, that grants added precision in two domains. First, overtime hours are commonly studied using either administrative surveys, or using some imputation rule (for example, labeling as overtime any hour above the 40th hour in a week): the Portuguese Personnel Records allow me to accurately classify any given worked hour in the data as either "regular" or overtime. Second, and more importantly, I do not have to make any assumptions about the worker's compensation scheme to infer the overtime pay, as the dataset directly includes the total overtime pay associated to the reported overtime hours.

Besides being able to track individuals and firms over the years, I am also able to identify the collective contract that covers each worker. While the data does not report the conditions of the contract itself, it does include, for each worker, a label that identifies the collective contract that applies to that worker. Because changes to the regulation of overtime compensation happen at the level of collective bargaining, it is particularly relevant that I am able to identify exactly this level in the data.

4.1.1 Sample Selection

I run my analysis on the universe of full-time entry level workers in retail and hospitality in Portugal, using data from 2006 to 2014.[23]

Table 1 shows how I identify workers who are in "entry-level jobs" in 2011: among *all* the workers in the 2006-2014 sample, 73% of the workers who enter the sample belong in one of the three bottom rungs of the firm's hierarchy (non-skilled, semi-skilled, and skilled), and average hourly wage only has a significant jump once we get to the fourth rung, i.e. higher-skilled. Therefore, I define as entry-level workers those who, in 2011,

[23]The reform I studied is introduced in 2012 and it applies until 2014, which is why I end my sample in that year.

are in one of the three lowest rungs of the ladder, excluding interns and trainees.[24] This corresponds to dropping around 27% of the individuals who are present in my data in 2011.

Jointly, retail and hospitality represent 28% of the observations, constituting the largest sector. I focus on this sector for two reasons. First, work is often organized in shifts (according to ONET data, 72% of respondents in retail say that they have an established routine and a set schedule, as opposed to one that changes with things like production demands),[25] which implies that overtime can be clearly defined as taking an extra shift. Indeed, the distribution of regular hours is very highly concentrated, with almost 80% of the observations bunching at the value of 40 hours per week. Second, output is fairly independent of other people's work: again using ONET data we can see that jobs in retail score high when asked about having little or no responsibility for work outcomes of other workers,[26] which implies that coordination across workers has limited importance and overtime work can be considered an individual decision.[27] Focusing on entry-level jobs is relevant for two reasons. First, the recorded overtime hours are likely to be inaccurate for executives, for whom the notion of overtime itself may be imprecise (Anxo and Karlsson, 2019). Second, these are probably jobs that entail routine tasks (salespersons in stores, or cleaning workers in hotels, are among the most common jobs in these rungs), in which the output is more likely to be proportional to the time spent working. Therefore, I expect the mechanism of signaling through overtime to matter more.[28]

The final sample includes the entire work histories from 2006 to 2014 of all workers in retail and hospitality who were in an entry-level job in 2011, and it comprises 406,222 individuals. Appendix D expands on the cleaning procedure, and it recaps the cleaning steps in a diagram that also shows how many observations are dropped at each stage, while

[24] Notice that I keep the entire working history of a worker, albeit truncated at 2006-2014, as long as they were in an entry-level job in 2011.

[25] Author's elaboration from ONET Release 20.1, occupations in retail are identified using ONET codes, prevalence of shift work is measured via the Work Context category "Work Schedule".

[26] Author's elaboration from ONET Release 20.1, need for coordination is measured via the Work Context category "Responsibility for Outcomes and Results".

[27] Also, for this sector there is no loss of generality in using data from the *Quadros de Pessoal*, and therefore in limiting the analysis to the private sector, which isn't true for sectors like health or energy where the public presence is relevant (Lopes and Tondini, 2024).

[28] Subjective measures of performance are more relevant in absence of verifiable output, like in cases where output isn't not necessarily proportional to hours, i.e. observable effort: "For some types of work (sales, for instance) [...] high-powered explicit incentives are optimal. For other types of work, however, there may be no good performance measures available at any reasonable cost. For employees doing these types of work, high-powered explicit incentive contracts are not optimal. [...] Career concerns, promotions and bonuses based on subjective performance evaluation, and employee selection (perhaps favoring those with high intrinsic motivation) are all alternative sources of incentives and motivation in circumstances where explicit incentive contracts are weak. (Baker, 2000)

highlighting the dimensions in which the dropped observations differ from the included ones.

4.1.2 Main Variables

The two main variables used in the analysis are the overtime premium and the promotion dummy.

The overtime premium is defined as the percentage difference between normal and overtime hourly wage. To compute it at the individual level, I follow Martins (2017) and I divide the base earnings of the month by the base hours of the month (to obtain the normal hourly wage), and the overtime earnings of the month by the overtime hours of the month (to obtain the overtime hourly wage).[29] The overtime premium for person i at time t is computed as:

$$\text{Overtime Premium}_{it} = \frac{\text{Overtime Hourly Wage}_{it} - \text{Normal Hourly Wage}_{it}}{\text{Normal Hourly Wage}_{it}} * 100.$$

So, for example, if during the normal workday the wage is 10 euros per hour, and overtime hours are paid time-and-a-half, 15 euros per hour, the overtime premium is 50%.[30] Notice that this measure is only defined for worker i and year t if she has performed at least one hour of overtime work in October of year t. This means that, for worker with zero overtime hours, the measure is missing, and it does not capture how much the worker would have been paid if she had done any overtime hours. To overcome this issue, I compute the mean overtime premium across all the workers covered by contract j and working in occupation r by averaging all the values of Overtime Premium$_{it}$ for which worker i in year t is covered by contract j and belongs to occupation r, and I attribute this measure to *all* the workers in that contract and that occupation, even if they didn't perform any overtime hours. An occupation here is defined by the "professional category": each collective bargaining contract defines different job titles for which potentially different conditions apply.[31] In this way, the overtime premium is defined for all workers such that at least one

[29] Since the data is collected in October for every year in my sample, seasonality is not an issue

[30] This definition is the closest I can get to the Labor Code quantities: the reason why this measure is not exactly the object of the law provisions is that the *Quadros de Pessoal* only record the total number of overtime hours the worker has performed in the reference month (and the corresponding payments), without indicating whether such overtime hours were performed right after the end of the workday, when the lowest overtime premium applies, or on a holiday, when the highest value of the range should be paid.

[31] Workers in a given "professional category", firm and year belong to a unique rung in the hierarchy in more than 97% of the cases. Conversely, workers who are in the same rung of the hierarchy, within a given firm and year, don't necessarily share the same professional category, as they may well have different tasks, roles, and contractual conditions.

of their "colleagues" has done overtime in October of year t. These represents 87.4% of the sample: Table D.3 in Appendix D shows how these observations differ from the estimation sample in the pre-reform years.[32]

The second key variable in my analysis is a dummy equal to 1 in year t if the worker is promoted by her employer that year. To construct it, I identify every time someone is recorded as being in a rung of the hierarchy higher than the one she was in the previous year,[33] while staying at the same firm. This last condition (not changing firms from one year to the next) is important because I am interested in studying promotions in relation to observed overtime hours: if a worker has just joined a firm, and in this move she climbs one rung of the ladder, previous overtime hours are probably unrelated to the carrier development, as they are unobserved by the current employer.[34] There are two versions of this variable: in one, the dummy is equal to 0 any time the worker has not climbed a rung and hasn't changed firm, or if the worker has changed firm (so that this variable is never missing); in another one, the variable is only equal to 0 if the worker has not climbed a rung, but it's instead set to missing if the worker is at different firm than the previous year (regardless of rung). Results included in the paper refer to the first definition, but my results are robust to using the second.

The fact that results using the two versions of the variable are consistent with each other is explained by the fact that there is relatively low mobility observed in my data: Table 2 reports the flows of workers from one year to the next, separately by rung in the ladder, and it shows that the probability of changing firm, from one year to next, is below 10% for all rungs of the hierarchy. The same table also helps understand why focusing on "internal promotions", and ignoring those that happen by changing employer, still depicts an accurate picture of workers' careers: the probability of changing firms while at the same time climbing one rung of the hierarchy is very low, and lower than 1% for most rungs. This is another important reason to focus on this context, and these sectors in particular: the mechanism I aim to identify is based on a private signal (overtime hours), and it is therefore particularly relevant in a setting where mobility across firms is relatively low.

Other relevant variables are overtime hours and overtime payments,[35], which, as men-

[32]This is the relevant measure of overtime premium I use to identify compliance with the law. In particular, given the institutional rules described in Section 3, I drop the observations for which overtime premium is "too inconsistent" with the Labor Code provisions, either by having a treatment variable smaller than 50% before 2012, or by having an increase in the overtime premium around 2012. These are 4.5% of the individuals in the full sample of entry-level workers from 2006-2014.

[33]64% of individuals who are in a higher level in the hierarchy, have climbed one rung only, so "rung skipping" is fairly common.

[34]In some cases this measure of promotion is associated to a decrease of normal hourly wage. Since these are only 6.4% of the tagged promotions, I simply recode these as 0's.

[35]The measure of overtime hours used in the analysis is winsorized at the 1% level, to account for fill in

tioned before, are recorded separately from normal hours and overtime hours, and pertain to the reference month in which the data is collected.

To measure worker quality, and therefore the selection the firm makes when promoting, I will consider tenure in the firm where the last promotion happened, total number of promotions, and average yearly wage growth in the worker career: tenure is reported directly in the data, and wage growth is computed as the percentage increase from the first year in the data, to the last one, divided by the number of years for which the worker is in the sample.

Appendix Table D.5 shows summary statistics for all the variables employed in the analysis. Appendix D includes additional descriptive tables of the sample. In particular, it shows how individuals who are excluded from the sample (either because the treatment variable is not defined because their contract-rung in 2011 has no worker doing overtime in any of the three pre-reform years, or because the treatment variable has some inconsistency with the law, or because they were not entry level in 2011) differ from the retained ones.

4.1.3 Descriptive Patterns

In this section I document the descriptive patterns that connect overtime hours, overtime premium, and promotions in the raw data. In particular, I test Prediction 1 and provide empirical evidence of the positive, concave relationship between overtime hours and promotions. I also show how this correlation becomes weaker as overtime premium increases, providing suggestive evidence in favor of Prediction 2.

While only 18% of the workers in my sample do overtime at least once in the sample years, the distribution of overtime hours has a long right tail: approximately three quarters of the workers that do at least one hour of overtime do 5 hours or more. Promotions are a rare event: on a given year 7.5% of the workers in my sample get promoted and only 8.6% of the workers in my sample get promoted twice or more in their career.[36] On average, workers who will be promoted in year $t+1$, do 1 hour of overtime in year t, while the same figure is equal to 0.82 for workers who won't be promoted the following year.

I now present two descriptive patterns that suggest the role of overtime as a signal within the firm. First, overtime hours are correlated with the probability of future promotion, and the shape of this relationship is consistent with the threshold structure derived in the model (so that "enough" overtime hours are needed, but increasing hours above a certain point does not further change future promotion probability). Second, such re-

[36] Average length of careers in my sample is 7.4 years (out of 9).

lationship is inversely related with overtime premium, consistently with the idea that for lower values of overtime pay, the signaling channel is stronger.

In order to study the relationship between overtime hours and future promotions more formally, I focus on the pre-reform sample, including only the years from 2006 to 2011, and I run the following regression:

$$Promoted_{i,t} = \alpha + \beta_1 Overtime\ Hours_{i,t-1} + \gamma \mathbf{X_{i,t}} + \delta_t + \eta_{j(i)} + \delta_t \tag{1}$$

where $Promoted_{i,t}$ is a dummy equal to 1 if individual i is, in year t, at a higher rung in the hierarchical ladder, relative to year $t-1$, and is in the same firm for years t and $t-1$. The variable $OvertimeHours_{i,t-1}$ represents a measure of overtime hours, for individual i in the year $t-1$, again conditional on i being in the same firm for the two consecutive years. The vector of controls $\mathbf{X_{i,t}}$ includes quadratic in age and tenure, a sex dummy, and the hourly wage for regular hours in year $t-1$, $\eta_{j(i)}$ are contract-by-occupation fixed effects, and δ_t are the year fixed effects. Standard errors are clustered at the contract-by-occupation level.

The estimates are reported in Table 3 for two different measures of $OvertimeHours_{i,t-1}$: column 1 uses a dummy variable equal to 1 for any amount of overtime hours, and column 2 uses a categorical variable that takes five values (0 overtime hours, 1 hour, 2-4 hours, 5-7 hours, 8 or more hours). The coefficient β_1 represents the correlation between overtime hours and future promotions: for both measures of overtime hours, this correlation is positive. While the magnitude of the estimate may look economically negligible, a comparison with the average probability of promotion reveals its importance: only 9.6% of workers that do not perform any overtime hours in year t are promoted the next year, and from column 1 we see that doing any overtime, for workers in the bottom tercile of the overtime premium distribution, increases the probability of promotion by 0.84 percentage points, which means an 8.75% increase of the baseline average promotion rate among the non-overtime workers.[37] While the positive correlation in column 1 is only significant at the 10% level, column 2 uncovers two important features of the structure of this relationship. First, the boost in future promotion probability is stronger if overtime hours are "enough", since doing 1 hour of overtime in year $t-1$ does not increase future promotion probability in year t, but doing 2-4 hours does. Second, further increases in the number of overtime hours do not lead to additional boosts in the probability of future promotions. This is consistent with the threshold structure of the equilibrium I derived: as explained in the discussion of Prediction 1, if each firm promotes workers who do "enough" over-

[37] This regression does not ascertain whether the uncovered correlation can be interpreted in a causal manner: we don't know if these workers are promoted *because* they did overtime they previous year, or if, on the contrary, they did overtime because they knew that doing so would grant them a promotion.

time hours, in the aggregate future promotion probability stops increasing as additional overtime hours are added.

Next, I study how the overtime premium impacts the relationship between overtime hours and future promotions

$$\begin{aligned} Promoted_{i,t} =& \alpha + \beta_1 Overtime\ Hours_{i,t-1} + \beta_2 Overtime\ Premium_{j(i),t-1} + \\ & \beta_3 Overtime\ Hours_{i,t-1} \times Overtime\ Premium_{j(i),t-1} + \gamma \mathbf{X_{i,t}} + \delta_t + \eta_{j(i)} + \delta_t \end{aligned} \tag{2}$$

where all the variables are defined as above and the overtime premium is computed as the average within the contract and occupation to which individual i belongs.[38] The results are presented in Table 4: the interaction between lagged overtime hours and lagged overtime premium are negative and significant, speaking to the intuition of Prediction 2. When overtime pay increases, the promotion boost associated to overtime hours decreases. For example, column 2 shows that, when the overtime premium is equal to 0, doing 2-4 hours of overtime increases future promotion probability by 9.55 percentage points, basically doubling the baseline promotion probability among non-overtime workers. Increasing the overtime premium by 100 percentage points (for example, as the overtime wage goes from being equal to normal wage to being double its size), reduced the effect of doing 2-4 overtime hours to 1.58 percentage points.

These two facts jointly represent suggestive evidence in favor of Predictions 1 and 2. The next section will describe the empirical strategy that exploits the variation provided by the 2012 reform.

4.1.4 Empirical Strategy

Identification Problem In principle, Predictions 2 and 3 could be studied in cross sectional data: they concern how the equilibrium of the signaling game changes at different levels of the overtime pay premium, and this variable displays a lot of variation in my data. However, an identification problem arises if the different levels of overtime pay premium are correlated with the promotion policy in a given firm. For example, it could be that firms that are facing negative demand shocks rely less on overtime hours, and at the same time are able to pay lower overtime wages: we would observe that for lower overtime premium there are lower overtime hours associated to promotions, for reasons that are completely spurious to the mechanism I am studying. The policy change described in section 3 provides a very useful setting since it is an arguably exogenous decrease in overtime

[38] The variable is rescaled by dividing it by 100, to facilitate interpretation of the magnitudes of the coefficients.

pay premium.

Treatment Definition My measure of overtime premium, as outlined in Section 4.1.2 is defined as the percentage difference between normal hourly wage and overtime hourly wage, averaged within a contract-by-occupation cell. To extract a measure of treatment intensity, I averaged the above measure for the three years before the reform (2009-2011). I attribute to each worker the value of treatment intensity of the contract-by-occupation she was in in 2011, the year before the reform.

As explained in Section 3, in the post-reform period, 2012 to 2014,[39] the overtime premium is fixed at the law-established level, as collective negotiation on this item is forbidden. Hence, the size of the drop in overtime premium depends on the pre-2012 level: higher overtime premium in 2011 means a bigger drop in overtime premium, hence higher treatment intensity. The actual size of the drop in overtime premium is not perfectly predicted by the pre-reform level, due to (potentially endogenous) non-compliance: I am therefore relying on the *de jure* variation, arguably exogenous.

Figure 4 shows the distribution of the treatment intensity variable, together with vertical dashed lines representing the terciles of this distribution. Figure 5 shows that the size of the drop in the overtime premium is positively correlated with the treatment intensity variable, confirming that my treatment intensity variable does indeed capture individuals' exposure to the reform.[40]

Identification Strategy My regressions will compare individuals with different levels of treatment intensity. Prediction 2 states that the overtime hours that are necessary for promotion decrease after the reform. To test it, I am going to verify that the reform leads to a decrease in overtime hours, particularly for workers who are about to be promoted, by estimating the following two regressions. First, I study whether the reform leads to a decrease in the overtime hours:

$$Overtime\ Hours_{it} = \beta_0 + \beta_1 Treatment\ Intensity_{i(j)} \times Post_t + \gamma \mathbf{X_{it}} + \eta_{j(i)} + \delta_t. \qquad (3)$$

Next, I study whether this decrease is observed also before promotion, besides the

[39]Law 23/2012 came into force on August 1st 2012, and collective bargaining agreements are prohibited to alter overtime premium until December 31st 2014. Since the reference month for the data collection of the Personnel Records is October, the year 2012 is already treated, and the year 2014 is still treated.

[40]In the regression tables included below, the treatment intesitity variable, whose range goes from 50 to 300, is divided by 100 to facilitate interpretation and discussion of the coefficients magnitude.

average response to the policy.

$$\begin{aligned} Overtime\ Hours_{i,t} = \beta_0 + \beta_1 Treatment\ Intensity_{j(i)} \times Post_t + \\ \beta_2 Promoted_{i,t+1} \times Post_t + \beta_3 Treatment\ Intensity_{j(i)} \times Post_t \times Promoted_{i,t+1} + \\ \gamma \mathbf{X_{i,t}} + \eta_{j(i)} + \delta_t, \end{aligned} \tag{4}$$

The identification assumption for Equation (3) entails that overtime hours would have changed in parallel for workers with different values of the treatment variable, absent the reform. For the parameter β_3 in Equation (4) to be identified, it must be the case that this parallel trend assumption holds regardless of the future promotion status of the worker. In particular this means that the promotion policy of the firm, and specifically how promotions are allocated to workers who supply different levels of overtime hours, would have changed in the same way regardless of the exposure to the reform.

The institutional setting supports this assumption: collective bargaining in Portugal rarely happens at the firm level, and contracts are often extended (either voluntarily or by government's mandate) to workers and firms that were not represented in the original negotiations (Addison et al., 2017; Martins, 2021). This speaks to the assumption that the contractual conditions are exogenous to the characteristics of the individual or the firm.[41] Furthermore, Appendix Table E.8 shows that firms who were most affected by the reform do not show different trends in terms of sales: this addresses concerns about potential confounding effects attributable to the 2010-2014 Portuguese financial crisis.[42] Further support for this assumption is provided by Figure 6, which shows the average number of overtime hours in the pre-reform period, separately for different levels of treatment intensity, defined by the terciles of the distribution of the treatment variable. From 2006 to 2011, the overtime hours changed in parallel for the different groups, i.e. for the different levels of treatment intensity.

Prediction 3 states that the average quality of promoted workers increases after the reform. To test it, I compare workers who are promoted before the reform (2009-2011) with workers who are promoted after the reform (2012-2014), hence excluding workers who are never promoted. In particular, I estimate the following regression:

[41] Addison et al. (2017) also mention that "collective agreements have tended to be revised regularly only insofar as wages are concerned, their other terms and conditions often being left untouched for many years". Overtime premium, therefore, is unlikely to be updated regularly, so that the 3-years average I use is exogenous to firm's growth, or year-to-year changes in demand and size.

[42] Notice that this would be a problem only if firms with different exposure to the reform were hit differently by the economic downturn.

$$
\begin{aligned}
Career_i = \beta_0 + \beta_1 Promoted\ Post_i + \beta_2 Treatment\ Intensity_{j(i)} + \\
\beta_3 Promoted\ Post_i \times Treatment\ Intensity_{j(i)} + \gamma \mathbf{X_i} + \eta_{j(i)},
\end{aligned}
\tag{5}
$$

where $Career_i$ is a time-invariant proxy of the worker's ability; $PromotedPost_i$ is a dummy equal to 1 if the worker is the worker's last promotion happened in the years 2012-2014 (post-reform period), and equal to 0 if the worker has only been promoted in the years 2009-2011 (pre-reform period), while it is missing for workers who are never promoted in their careers, so that the coefficient can be interpreted as the differential effect of being promoted after 2012 versus being promoted before. This excludes the confounding effect of comparing the careers of promoted workers with that of non-promoted workers. All the other variables have the same interpretation as in the previous specifications, but notice that the vector of controls only include time-invariant variables (age in 2011, normal hourly wage in 2011, a sex dummy). Notice that in principle workers with $Promoted\ Post_i = 0$ could be negatively selected, as they are defined as those who where not promoted again after one promotion in the 2009-2011 period. However, an alternative specification that includes as control the rung to which the worker belonged when she was promoted last shows very similar results. This point mitigates such concerns, as the comparison is done only among people who were in the same level of the hierarchy when they were promoted last.

Identification of the coefficient of interest, β_3 entails assuming that the timing of promotion would not have had any impact on the careers of workers with different exposure to the reform, absent the reform itself. The considerations discussed before, and in particular the observation that differentially affected firms did not show different performances over the sample years, lend support in favor of this assumption.

5 Results

This section presents the main empirical results of the paper. I show that following the 2012 reform that lowered the overtime premium, the most affected workers decrease their overtime hours more, in particular if they are about to be promoted, and are better selected if promoted. In particular, Section 5.1 tests prediction 2 by showing that overtime hours decrease, and in particular they do so before promotion. Section 5.2 tests prediction 3 by showing that the reform increases the quality of promoted people. Appendix E shows the effect of the reform on some secondary variables, estimated by varying the outcome variable in 3. These results confirm that there was no compensation to normal hourly wage, and no effects on the probability of changing firm.

5.1 Results on Overtime Hours

The first implication of Proposition 2 is that the threshold of overtime hours required for promotion decreases. Observationally, this implies that overtime hours decrease, and in particular that they decrease for workers who are about to be promoted. The estimates of Equation (3) and (4) confirm that that's the case.

Column 1 of Table 5 shows that the reform leads to a decrease in overtime hours: if treatment exposure increases by 100 (e.g. the pre-reform overtime wage goes from being equal to the normal hourly wage to being twice of it), workers do 0.32 overtime hours less, which is 40% of the baseline average of overtime hours.

Table 6 shows how this average decrease affects workers who are about to be promoted as well: we can see that workers who are about to be promoted do more overtime hours compared to the other workers in the post period as well (coefficient β_2 of Equation (4) is positive), but they decrease their overtime hours compared to their pre-reform selves. The magnitude of this effect is sizeable: if treatment exposure increases by 100 (e.g. the pre-reform overtime wage goes from being equal to the normal hourly wage to being twice of it), workers who are about to be promoted do 0.56 overtime hours less, which is 70% of the pre-reform average among non-promoted workers.

This result speaks to the fact that overtime hours are even more conducive to future promotions, for the most treated workers, in the post period, which is precisely the intuition of Proposition 2.

5.2 Results on Quality of Promoted Workers

The second implication of Proposition 2 is that the average quality of promoted people increases (Prediction 3).

Table 7 shows the higher exposure to the reform determines, for workers who are promoted at least once after 2012, better selection relative to those whose last promotion was before 2011, and relative to the workers less exposed to the reform. The different outcomes used to proxy for quality of the selection (suitability to the firm) are the total number of promotions, the average yearly wage growth over one's career (both raw and net of the baseline wage growth of the economy), and tenure in the firm in which the worker was when she was promoted for the last time. While the latter may not represent quality of the worker *per se* (one may argue that the best workers are more likely to be poached) it does represent, from the firm's perspective, a desirable outcome. In other words, the firm is doing a better job at selecting workers if it's able to retain the promoted ones.

While on average workers who are promoted at least once after the reform do not seem positively selected (and if anything, they have a lower average yearly wage growth), the diff-in-diff coefficients all show that exposure to the reform leads to increased quality of promoted workers. In particular, an increase in treatment intensity of 100 percentage points leads to almost half a promotion more over the course of the career of workers who are promoted after the reform (a 33% increase relative to the average of 1.3 promotions for the workers whose last promotion is before the reform), to an increase of 0.85 percentage points in the average annual wage growth rate (26% increase relative to the baseline mean), and to 1.6 years longer tenure in the firm where the last promotion happened (18% of the baseline mean).

Overall, these results confirm that the prediction of Proposition 2: a decrease in the overtime pay premium leads to an increase in the quality of promoted people.

Alternative theories of promotion This set of results is crucial for the overall interpretation of my empirical analysis: this section discusses how the observed increase in the quality of promoted people (jointly with the observed decrease in overtime hours documented in section 5.1) is inconsistent with alternative leading theories of internal labor market, like human capital accumulation and incentive provision, as sole explanations of the observed relationship between overtime hours and promotions outlined in section 4.1.3.

The intuitive relationship between current effort and future promotions can be rationalized under different models. In particular, human capital accumulation and incentive provision would both generate the same relationship between overtime hours and future promotion that I observe in the cross sectional data.[43]

I argue that under these alternative theoretical frameworks a decrease in the overtime pay premium would generate different predictions. This rules out each of these alternative stories as unique explanations of the role of overtime hours on future promotion probability. First, human capital accumulation implies that workers learn while doing overtime, and promotion reflects their improved ability. My results however show that overtime hours decrease (in particular, before promotions) *and* average quality of promoted people

[43] For a review of the theoretical and empirical literature that studies promotions as incentives, or motivational tools to encourage workers to exert effort, see Georgiadis (Forthcoming). The idea that promotions follow to human capital accumulation dates back to Becker (1962). See also, for example, Prendergast (1993) and Kwon and Milgrom (2014).

increases, a pattern that cannot be explained by the sole association of overtime hours and learning. Second, incentive provision implies that overtime hours are beneficial for the firm and the promotion prize associated to them substitutes the increased hourly wage workers receive when doing overtime. Under this view, the reform would lead firm to promote overtime workers more, in order to maintain the incentives for their employees to supply overtime hours even at the reduced overtime wage. Once again, the results on the quality of promoted people rule out this as the unique explanation of promotions: if firms had increased the promotion probability associated to overtime hours in response to the reform, they would select worse workers, not better ones.

6 Conclusions

In this paper, I have developed and tested a theory of signaling within the firm, which is a context where a part of the cost of signaling is bore by the receiver (i.e. the firm) in the form of compensation to the worker. In the context of my analysis, this signal is overtime work, which is paid at a higher rate than work performed during normal business hours, and which is at same time costly for workers to supply. In my model, the cost of working overtime hours is negatively correlated with unobserved worker characteristics, which in turn is a measure of the worker contribution to the firm's profit once promoted. Under this setting, the firm wants to use overtime hours as a signal to decide which workers to promote. In my model, overtime premium and wages are exogenous, consistently with the empirical setting in which I test its prediction. This means that the only tool the firm has to control selection is the relationship between overtime hours and promotion probability. Under these assumptions, three predictions follow. First, the firm's optimal promotion policy is to promote workers whose overtime hours surpass a certain threshold, and only them. Second, an exogenous decrease in the overtime premium implies a decrease of this threshold. Third, this policy implies increased separation in the equilibrium choice of overtime hours, and therefore an increase in the quality of promoted people.

Next, I bring these predictions to the data. First, I show a cross-sectional relationship between overtime hours and future promotion, that fades away as overtime hours further increase, consistently with a threshold structure of the promotion policy. To test the remaining predictions, I provide causal evidence by exploiting a policy change in Portugal that lowered overtime premium from 2012 to 2014. Since exposure to the reform was heterogeneous across different collective bargaining contracts, I can implement a diff-in-diff design and show that the most exposed workers (those for which overtime premium decreased the most) work less overtime hours on average, particularly before promotion, indicating that fewer overtime hours are sufficient to be promoted. Conditional on being promoted at least once in the years 2009-2014, workers who are promoted after the reform are better selected according to a battery of proxies.

These results represent the first quasi experimental evidence of signaling within the firm, and jointly rule out human capital accumulation or incentive provision as the sole explanation of the positive correlation between overtime hours and promotions.

More broadly, these results inform two sets of considerations. First, from the firm's perspective, subordinating promotion to an observable measure of effort, that is also independently rewarded, means incurring in some "wrong" promotion decisions, given the pooling nature of the unique equilibrium that my model generates. Thinking about "optimal promotion policies" requires taking into account the strategic response workers will have to such policies. Second, my model has assumed that the cost of supplying overtime hours is negatively related to the unobserved suitability to the firm, in a deterministic way. In other words, the relationship between the worker's type and her cost of working long hours is the same for all workers. Conversely, it seems reasonable to believe that such relationship may be heterogeneous, and in particular that some workers may be unable to supply long working hours, regardless of their type. If the firms does not correctly take into account such heterogeneity when making promotion decisions, then some talent will be left out of successful careers.

References

ACEMOGLU, D., D. AUTHOR, O. , ASHENFELTER, ET AL. (2011): "Skills, tasks and technologies: Implications for employment and earnings," *Handbook of Labor Economics*, 4, 1044–1171.

ADDISON, J. T., P. PORTUGAL, AND H. VILARES (2017): "Unions and collective bargaining in the wake of the Great Recession: evidence from Portugal," *British Journal of Industrial Relations*, 55, 551–576.

ALTONJI, J. G. AND C. R. PIERRET (2001): "Employer learning and statistical discrimination," *The Quarterly Journal of Economics*, 116, 313–350.

ANGER, S. (2008): "Overtime work as a signaling device," *Scottish Journal of Political Economy*, 55, 167–189.

ANXO, D. AND M. KARLSSON (2019): *Overtime work: A review of literature and initial empirical analysis*, International Labour Office.

BAKER, G. (2000): "The use of performance measures in incentive contracting," *American Economic Review*, 90, 415–420.

BARLEVY, G. AND D. NEAL (2019): "Allocating effort and talent in professional labor markets," *Journal of Labor Economics*, 37, 187–246.

BECKER, G. S. (1962): "Investment in human capital: A theoretical analysis," *Journal of political economy*, 70, 9–49.

BEDARD, K. (2001): "Human capital versus signaling models: university access and high school dropouts," *Journal of political economy*, 109, 749–775.

BELL, L. A. AND R. B. FREEMAN (2001): "The incentive for working hard: explaining hours worked differences in the US and Germany," *Labour Economics*, 8, 181–202.

BENSON, A., D. LI, AND K. SHUE (2019): "Promotions and the Peter principle," *The Quarterly Journal of Economics*, 134, 2085–2134.

BOGNANNO, M. AND E. MELERO (2016): "Promotion signals, experience, and education," *Journal of Economics & Management Strategy*, 25, 111–132.

BOSSLER, M. AND P. GRUNAU (2020): "Asymmetric information in external versus internal promotions," *Empirical Economics*, 59, 2977–2998.

BRATTI, M. AND S. STAFFOLANI (2007): "Effort-based career opportunities and working time," *International Journal of Manpower*, 28, 489–512.

CARDOSO, A. R. AND P. PORTUGAL (2005): "Contractual wages and the wage cushion under different bargaining settings," *Journal of Labor Economics*, 23, 875–902.

CARVALHO, C. D. O. AND A. T. RIBEIRO (2022): "The impact of the EU economic governance in Portugal," *European Labour Law Journal*, 13, 193–213.

CHO, I.-K. AND D. M. KREPS (1987): "Signaling games and stable equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

CLARK, D. AND P. MARTORELL (2014): "The signaling value of a high school diploma," *Journal of Political Economy*, 122, 282–318.

ESTEBAN, J. AND D. RAY (2006): "Inequality, lobbying, and resource allocation," *American Economic Review*, 96, 257–279.

FARBER, H. S. AND R. GIBBONS (1996): "Learning and wage dynamics," *The Quarterly Journal of Economics*, 111, 1007–1047.

FREDERIKSEN, A., T. KATO, AND N. SMITH (2024): "Working Hours, Top Management Appointments, and Gender: Evidence From Linked Employer-Employee Data," *Journal of Labor Economics*.

GEORGIADIS, G. (Forthcoming): "Contracting with moral hazard: A review of theory and empirics," in *Elgar Encyclopedia on the Economics of Competition and Regulation*, Michael Noel.

GHOSH, S. AND M. WALDMAN (2010): "Standard promotion practices versus up-or-out contracts," *The RAND Journal of Economics*, 41, 301–325.

GIBBONS, R. AND L. F. KATZ (1991): "Layoffs and lemons," *Journal of labor Economics*, 9, 351–380.

GIBBONS, R. AND M. WALDMAN (1999): "A theory of wage and promotion dynamics inside firms," *The Quarterly Journal of Economics*, 114, 1321–1358.

HEYWOOD, J. S. AND X. WEI (2004): "Education and signaling: evidence from a highly competitive labor market," *Education Economics*, 12, 1–16.

HUITFELDT, I., A. R. KOSTØL, J. NIMCZIK, AND A. WEBER (2023): "Internal labor markets: A worker flow approach," *Journal of Econometrics*, 233, 661–688.

HUNNES, A. (2012): "Testing the role of comparative advantage and learning in wage and promotion dynamics," *International Journal of Manpower*, 33, 556–582.

INTERNATIONAL MONETARY FUND (2011): "Portugal: Letter of Intent, Memorandum of Economic and Financial Policies, and Technical Memorandum of Understanding," .

KAHN, L. B. (2013): "Asymmetric information between employers," *American Economic Journal: Applied Economics*, 5, 165–205.

KAHN, L. B. AND F. LANGE (2014): "Employer learning, productivity, and the earnings distribution: Evidence from performance measures," *The Review of Economic Studies*, 81, 1575–1613.

KWON, I. AND E. M. M. MILGROM (2014): "The significance of firm and occupation specific human capital for hiring and promotions," *Labour Economics*, 31, 162–173.

LANDERS, R. M., J. B. REBITZER, AND L. J. TAYLOR (1996): "Rat race redux: Adverse selection in the determination of work hours in law firms," *The American Economic Review*, 329–348.

LANG, K. AND D. KROPP (1986): "Human capital versus sorting: the effects of compulsory attendance laws," *The Quarterly Journal of Economics*, 101, 609–624.

LAZEAR, E. P. AND P. OYER (2004): "The structure of wages and internal mobility," *American Economic Review*, 94, 212–216.

LIMA, F. AND P. TELHADO PEREIRA (2003): "Careers and wages within large firms: evidence from a matched employer-employee data set," *International Journal of Manpower*, 24, 812–835.

LLUIS, S. (2005): "The role of comparative advantage and learning in wage dynamics and intrafirm mobility: Evidence from Germany," *Journal of Labor Economics*, 23, 725–767.

LOPES, M. C. AND A. TONDINI (2024): "Firm-Level Effects of Reductions in Working Hours," *mimeo*.

MARKUSSEN, S. (2009): "The effects of sick-leaves on erarnings," Tech. rep., Memorandum.

MARTINS, P. S. (2017): "Economic effects of overtime premium flexibility: Firm-and worker-level evidence from a law reform," Tech. rep., GLO Discussion Paper.

——— (2021): "30,000 minimum wages: The economic effects of collective bargaining extensions," *British Journal of Industrial Relations*, 59, 335–369.

MEYER, A. AND M. WALLETTE (2005): "Absence of Absenteeism and Overtime Work–Signalling Factors for Temporary Workers?" Tech. rep., Working Paper.

PANNENBERG, M. (2005): "Long-term effects of unpaid overtime: evidence for West Germany," *Scottish Journal of Political Economy*, 52, 177–193.

PASTORINO, E. (2024): "Careers in firms: the role of learning about ability and human capital acquisition," *Journal of Political Economy*, 132, 1994â2073.

PINKSTON, J. C. (2009): "A model of asymmetric employer learning with testable implications," *The Review of Economic Studies*, 76, 367–394.

PRENDERGAST, C. (1993): "The role of promotion in inducing specific human capital acquisition," *The Quarterly Journal of Economics*, 108, 523–534.

ROSEN, S. (1985): "Prizes and incentives in elimination tournaments," .

SCHÖTTNER, A. AND V. THIELE (2010): "Promotion tournaments and individual performance pay," *Journal of Economics & Management Strategy*, 19, 699–731.

SPENCE, M. (1973): "Job Market Signalling," *The Quarterly Journal of Economics*, 87, 355–374.

TÔ, L. T. (2018): "The signaling role of parental leave," *Mimeo*.

TREBLE, J., E. VAN GAMEREN, S. BRIDGES, AND T. BARMBY (2001): "The internal economics of the firm: further evidence from personnel data," *Labour Economics*, 8, 531–552.

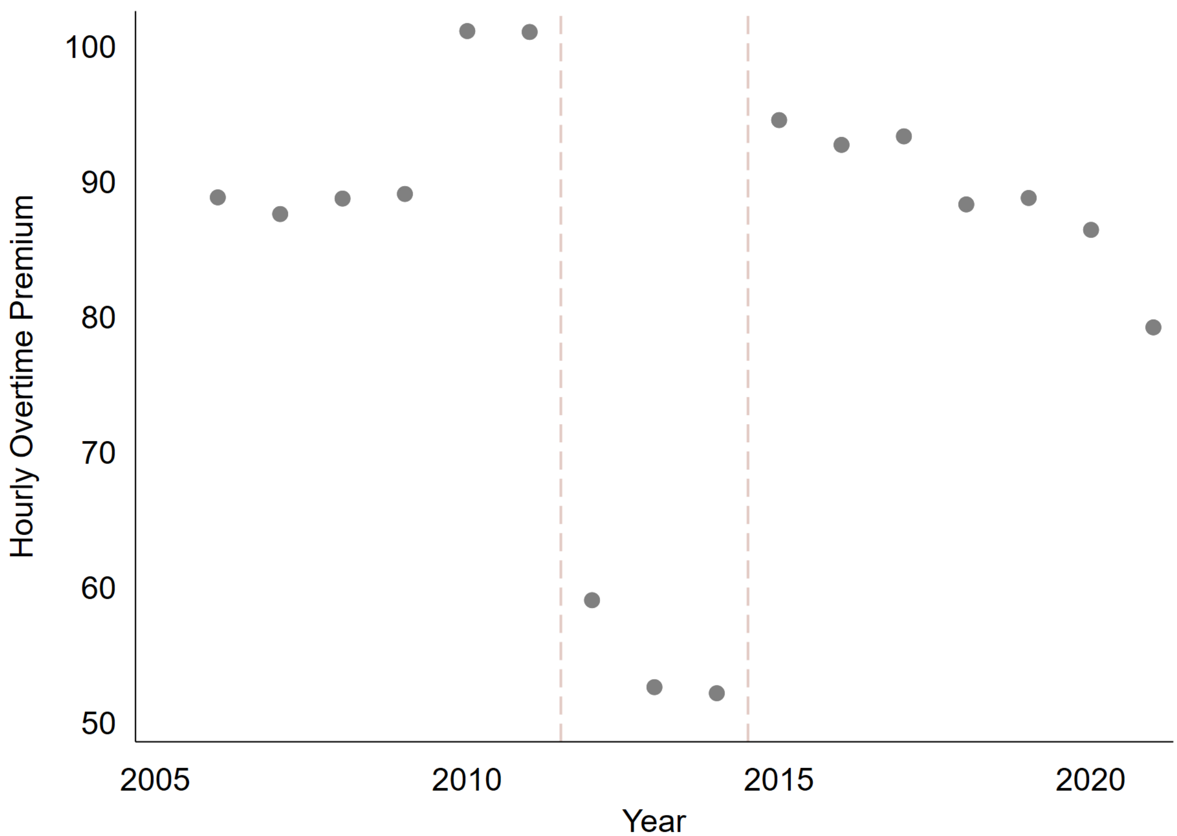
VAN DER KLAAUW, B. AND A. DIAS DA SILVA (2011): "Wage dynamics and promotions inside and between firms," *Journal of Population Economics*, 24, 1513–1548.

WALDMAN, M. AND Z. YIN (2024): "Promotions, adverse selection, and efficiency," *Journal of Labor Economics*.

WU, A. H. (2024): "Reveal or Conceal? Employer Learning in the Labor Market for Computer Scientists," Tech. rep., Working Paper.

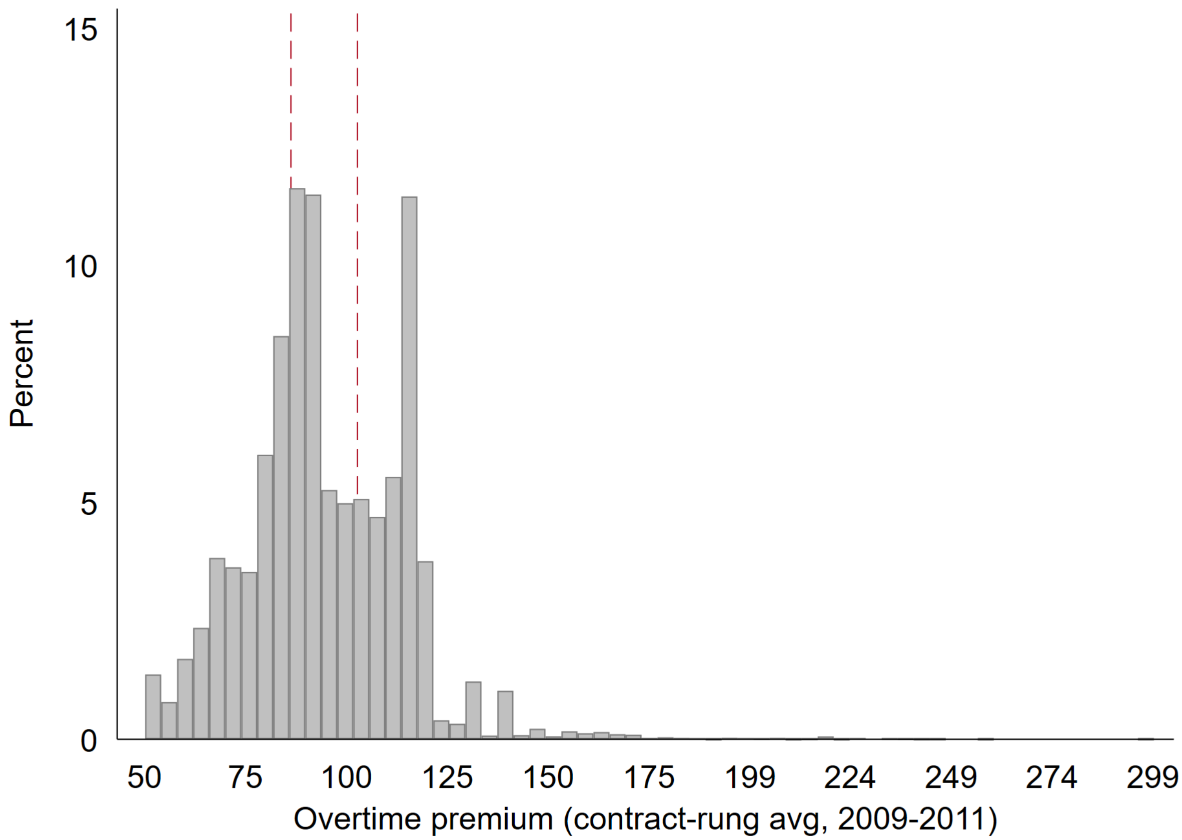
7 Figures

Figure 3: Time evolution of average overtime premium



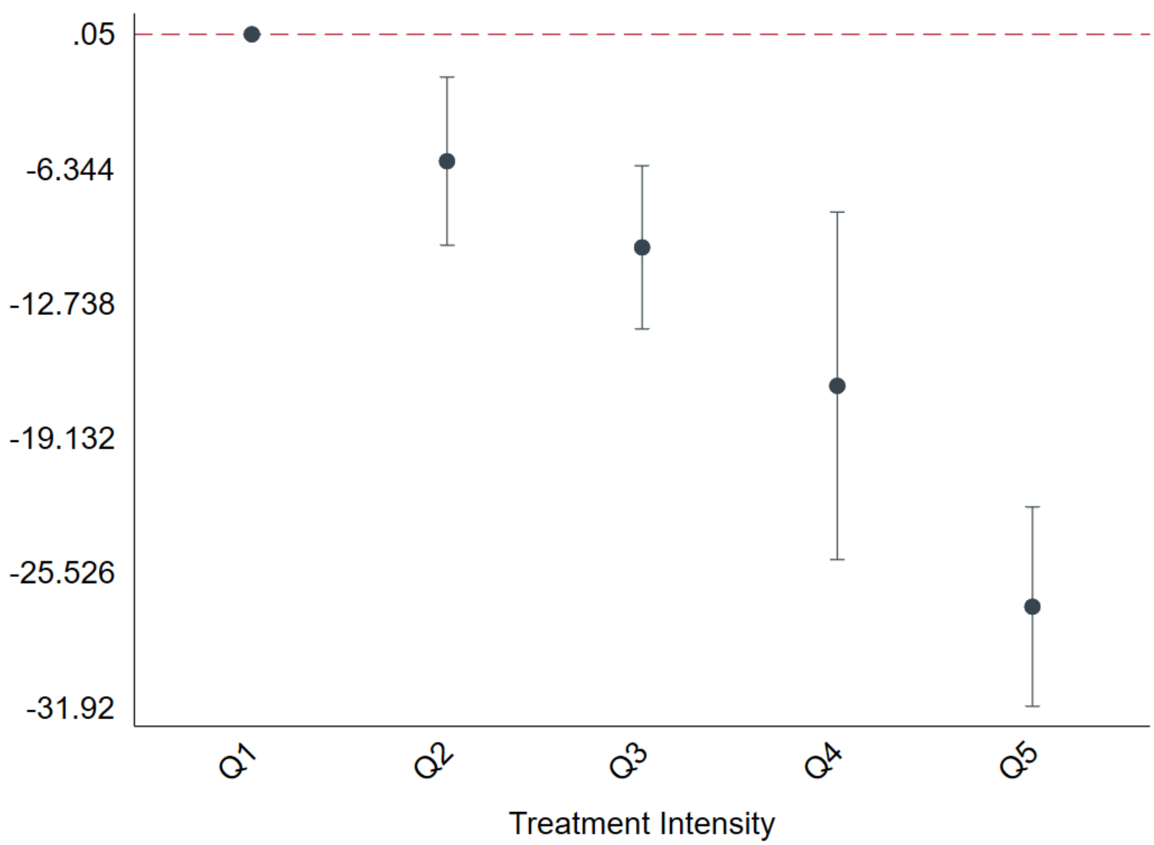
Note: This figure shows the time evolution of the average overtime premium, defined as the percentage difference between hourly pay for overtime hours, compared to normal hourly pay. The first vertical dashed line represents the introduction of the law, which lowered overtime premium by 50% on average, and enforced the new value for two years. The second vertical dashed line represents the end of the enforcement, i.e. the moment when collective bargaining was allowed again to modify the premium established in the Labor Law.

Figure 4: Treatment variable distribution



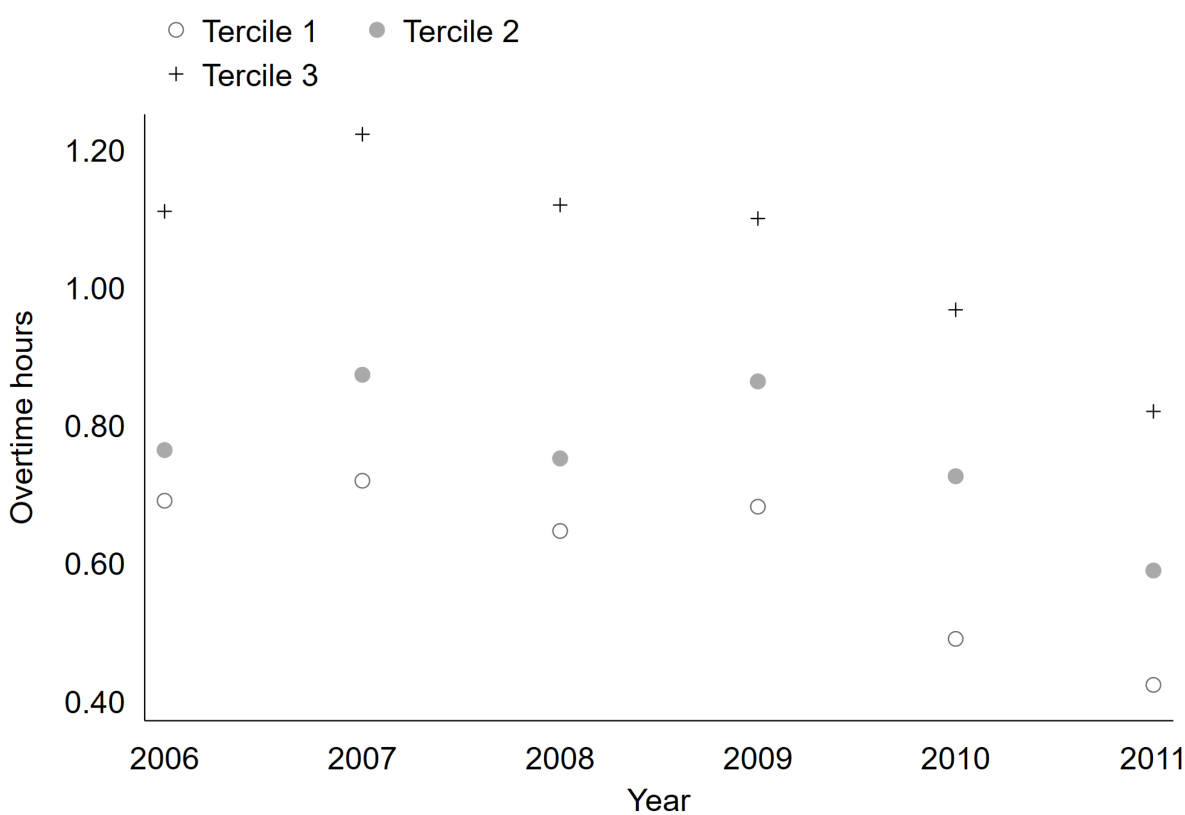
Note: This figure represents the distribution of the variable that represents treatment intensity: the average overtime premium in a contract-by-occupation cell, on average in the years 2009-2011. Each individual is attributed the value of the variable for the contract and occupation where she was employed in 2011. The vertical dashed lines represent the terciles of the distribution (below 86%, from 86% to 102%, above 102%).

Figure 5: First stage: Overtime Premium decreases more as Treatment Intensity increases



Note: This figure shows the results of the "first stage" regression that verifies whether the workers with higher values of the treatment intensity variable experience indeed a bigger drop in the overtime premium. The specification is $Overtime\ Premium_{it} = \beta_0 + \beta_1 Post_t \times + Treatment\ Intensity_{j(i)} \mathbf{X_{it}} + \eta_{j(i)} + \delta_t$, where $Overtime\ Premium_{it}$ is the percentage difference between normal hourly wage and overtime hourly wage the individual i receives at time t, $Post_t$ is a dummy equal to 1 from 2012 onwards, $\eta_{j(i)}$ are contract-by-occupation fixed effects, and δ_t are year fixed effects. The variable $TreatmentIntensity_{j(i)}$ is the contract-by-occupation average overtime premium, on average from 2009 to 2011. Instead of being used as a continuous variable, here it is replaced by a categorical variable, whose five values correspond to quintiles of is distribution (below 80%, 80-88%, 88-97%, 97-113%, above 113%). The equivalent regression estimated using the continuous measure of treatment intensity is reported in column 2 of Table 5.

Figure 6: Parallel trends in overtime hours before the reform



Note: This figure shows the average number of overtime hours in the pre-reform years, separately for different levels of treatment intensity, identified by the terciles of the distribution plotted in Figure 4.

8 Tables

Table 1: Summary statistics: careers in retail and hospitality

Hierarchy rung	% entrants	Avg hourly wage (base)	Avg over-time hours	Probability of promotion	1-yr wage growth if promoted	3-yr wage growth if promoted	5-yr wage growth if promoted	Base earnings
Top management	0.04	11.71	0.72	0.00	.	.	.	1988.07
Middle management	0.03	8.12	0.28	0.04	6.15	14.24	28.47	1381.51
Supervisors	0.03	6.53	0.78	0.03	7.71	15.88	28.87	1119.19
Higher-skilled	0.06	6.23	1.35	0.04	7.46	13.62	26.77	1062.28
Skilled	0.38	3.87	0.73	0.02	9.48	17.33	29.46	660.16
Semi-skilled	0.20	3.26	0.97	0.08	6.78	12.25	23.28	544.95
Non-skilled	0.15	3.02	0.87	0.16	6.68	13.07	27.86	478.08
Interns, trainees	0.11	2.90	0.87	0.21	8.40	16.69	31.63	479.86

Note: This table reports summary statistics about worker careers, for the different rungs of the hierarchy ladder. The column "% of entrants" reports how many, of the workers who are in the data for the first time, are in a given rung. The rest of the column titles are self-explanatory. The sample is the universe of workers in retail and hospitality, from 2006 to 2014.

Table 2: Transition probabilities by hierarchy rungs

Hierarchy rung		% Same firm, not promoted	% Same firm, promoted	% Other firm, not promoted	% Other firm, promoted	% Not in sample
Top management	4056	73.62	.	8.33	.	18.05
Middle management	5147	68.74	1.77	8.02	0.17	21.29
Supervisors	10819	69.80	1.33	9.51	0.26	19.10
Higher-skilled	24544	72.68	1.44	8.55	0.30	17.03
Skilled	1329531	81.53	0.93	3.70	0.16	13.67
Semi-skilled	502797	73.09	6.18	2.93	1.13	16.66
Non-skilled	216134	62.04	13.76	2.33	2.16	19.72
Interns, trainees	48626	38.06	31.69	1.93	9.02	19.30

Note: This table reports the transition probabilities, from one year to the next, across different states, separately for the different rungs in the hierarchy.

Table 3: Correlation between overtime hours and future promotions

	(1) Promoted$_t$	(2) Promoted$_t$
Any Overtime$_{t-1}$	0.00840	
	(0.00514)	
Overtime Hours$_{t-1}$		
1		0.0092
		(0.0108)
2-4		0.0263*
		(0.0124)
5-7		0.0117
		(0.0116)
8+		0.0031
		(0.0033)
Observations	916,988	916,988
Unique IDs	310,999	310,999
Year FEs	Yes	Yes
Contract × Occupation FEs	Yes	Yes
Mean Promoted$_t$, for Overtime Hours$_{t-1}$ = 0	0.0963	0.0963

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports the results from a regression of the promotion dummy on different measured of lagged overtime hours, controlling for quadratic in tenure, quadratic in age, normal hourly wage (lagged), and female dummy. The regression includes year and contract-by-occupation fixed effects and standard errors are clustered at the contract-by-occupation level. The reported mean for the dependent variable is computed on the sample of workers who have 0 overtime hours. The regression is run on the 2006-2011 sample only.

Table 4: Overtime pay is inversely related to the correlation between overtime hours and future promotions

	(1) Promoted$_t$	(2) Promoted$_t$
Overtime Premium $_{t-1}$	-0.000388**	-0.0386**
	(0.000126)	(0.00016)
Any Overtime$_{t-1}$	0.0324***	
	(0.0125)	
Any Overtime$_{t-1}$ × Overtime Premium $_{t-1}$	-0.000294**	
	(0.0000945)	
Overtime Hours$_{t-1}$		
1		0.164**
		(0.0596)
2-4		0.0955**
		(0.0345)
5-7		0.0160
		(0.0166)
8 +		0.0221
		(0.0121)
Overtime Hours$_{t-1}$ × Overtime Premium $_{t-1}$		
1		-0.154**
		(0.0564)
2-4		-0.0797*
		(0.0315)
5-7		-0.0105
		(0.0175)
8 +		-0.0228
		(0.0125)
Observations	916,068	916,068
Unique IDs	310,938	310,938
Year FEs	Yes	Yes
Contract × Ocupation FEs	Yes	Yes
Mean dep. var	0.0963	0.0963

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table reports the results from a regression of the promotion dummy on different measured of lagged overtime hours, interacted with lagged overtime premium (averaged within contract-by-occupation cells), and controlling for quadratic in tenure, quadratic in age, normal hourly wage (lagged), and female dummy. The regression includes year and contract-by-occupation fixed effects and standard errors are clustered at the contract-by-occupation level. The reported mean for the dependent variable is computed on the sample of workers who have 0 overtime hours. The regression is run on the 2006-2011 sample only.

Table 5: Effects of overtime reform: baseline outcomes

	(1) Overtime hours$_{i,t}$	(2) Hourly overtime premium$_{i,t}$
Post$_t \times$ Treatment Intensity$_{j(i)}$	-0.319***	-58.74***
	(0.0936)	(3.996)
Observations	2,348,005	157,910
Unique IDs	398,546	72,795
Year FEs	Yes	Yes
Contract $\times$ Occupation FEs	Yes	Yes
Mean dependent variable (Post$_t$ = 0)	0.768	96.68

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the effect of the 2012 reform on Overtime Hours and Hourly Overtime Premium, estimated according to the following estimation Equation:

$$Outcome_{it} = \beta_0 + \beta_1 Treatment\ Intensity_{i(j)} \times Post_t + \gamma \mathbf{X_{it}} + \eta_{j(i)} + \delta_t,$$

where the vector of controls includes quadratic in age and tenure, hourly wage for normal hours, and a sex dummy. The regression included contract-by-occupation fixed effects $\eta_{j(i)}$ and year fixed effects δ_t. Standard errors are clustered at the contract-by-occupation level.

Table 6: Effects of overtime reform on promotion policy

	(3) Overtime Hours$_{i,t}$
Post$_t$ × Treatment Intensity$_{j}(i)$	-0.340**
	(0.108)
Promoted$_{i,t+1}$ × Treatment Intensity$_{j}(i)$	-0.0463
	(0.138)
Promoted$_{i,t+1}$ × Post$_t$	0.500*
	(0.212)
Promoted$_{i,t+1}$ × Post$_t$ × Treatment Intensity$_{j}(i)$	-0.559*
	(0.241)
Observations	1,700,613
Unique IDs	365,858
Year FEs	Yes
Contract-by-rung FEs	Yes
Mean dep. var	0.757

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the effects of the 2012 reform on the association between overtime hours and future promotions, estimated according to Equation (4). The reported mean for the dependent variable is computed on the pre-reform sample (2006-2011, included), and only on workers who were not promoted in year t.

Table 7: Effects of overtime reform on selection

	(1) No. promotions	(2) Avg yearly wage growth	(3) Net avg yearly wage growth	(4) Tenure
Promoted Post$_i$	-0.139	-0.841**	-16.09**	-1.026*
	(0.104)	(0.306)	(5.591)	(0.435)
Treatment Intensity$_{j(i)}\times$ Promoted Post$_i$	0.446***	0.887*	14.90*	1.835***
	(0.126)	(0.363)	(6.723)	(0.493)
Observations	91,262	91,262	91,262	91,262
Contract × Occupation Fe	Yes	Yes	Yes	Yes
Mean dependent variable (Promoted Post$_i$ = 0)	1.38	3.203	53.71	8.793

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table represented the effects of the 2012 reform on quality of promoted people, estimated according to Equation (5). “Number of promotions” is computed as the total number of promotions over one’s career. “Average yearly wage growth” is computed as the total wage growth from the entrance in the sample to 2014, divided by the number of years the worker is in the sample. “Net average yearly wage growth” is computed as follows. First, I compute the overall annual wage growth across all workers in my sample from any given year y to year 2014. Next, for a worker who enters the sample in year y_1 and leaves it in year y_2, I define her average yearly wage growth between these two years, and then I divide it by the baseline average yearly wage growth for people who entered the sample in year y_1. “Tenure” is measured in years and refers to the firm in which the worker’s last promotion happened. The reported mean for the dependent variable is computed on the pre-reform sample (2006-2011, included), and only on workers whose last promotion happened before the reform (2009-2011).

A Proof of Proposition 1

The proof of Proposition 1 proceeds in several steps. In steps 1-4, I argue that in any equilibrium of the unrefined game there is an interval structure, where some types perform their myopic first best and some other pool at given values of h. In steps 5 and 6 I argue that every refined equilibrium has exactly one pooling interval, and in steps 7-9 I characterize this interval by identifying its extreme points.

Step 1

In any equilibrium, the promotion decision is

$$\pi(h) \begin{cases} = 1 & \text{if } \mathbb{E}(\theta \mid h) > 0 \\ \in [0,1] & \text{if } \mathbb{E}(\theta \mid h) = 0 \\ = 0 & \text{if } \mathbb{E}(\theta \mid h) < 0 \end{cases}$$

This directly follows the definition of equilibrium, for which the firm's promotion decision maximizes the firm's expected profits in the second period. Such expected profits are positive if and and only if the expected value of θ is positive. If $\mathbb{E}(\theta \mid h) = 0$, the firm is indifferent, and therefore any promotion rule satisfied the equilibrium condition. Notice that th expectation $\mathbb{E}(\theta \mid h)$ is computed with respect to the equilibrium posterior beliefs.

Step 2

In any equilibrium, the overtime choice $h(\theta)$ is non-decreasing.

To see this, consider two types $\theta 1, \theta_2$ such that $\theta_1 < \theta_2$, and let $h_i = h(\theta_i)$ for $i = 1, 2$. Since each worker is choosing her best response, none of them wants to mimic the other, i.e.

$$wh_1 - c(h_1, \theta_1) + \pi(h_1)V \geq wh_2 - c(h_2, \theta_1) + \pi(h_2)V,$$

and

$$wh_2 - c(h_2, \theta_2) + \pi(h_2)V \geq wh_1 - c(h_1, \theta_2) + \pi(h_1)V.$$

By summing up the two conditions:

$$c(h_1, \theta_2) - c(h_2, \theta_2) \geq c(h_1, \theta_1) - c(h_2, \theta_1)$$

Say, by way of contradiction, that $h_1 > h_2$. Then both sides of the above expression will be negative as the cost increases with overtime hours. Therefore, the expression holds if the the left-hand side is the one closer to zero, i.e. if type θ_2 is the one for which the cost increases the most when overtime hours increases. This contradicts my assumption on the cost function for which the $\frac{\partial^2}{\partial\theta\partial h}c(\cdot,\theta) < 0$, meaning that higher types have smaller changes in cost when overtime hours increase. I conclude that $h_1 \leq h_2$.

Step 3

Define a pool to be any maximal interval $[\theta_1, \theta_2]$ with $\theta_1 < \theta_2$ such that $h(\theta) = h \forall \theta \in [\theta_1, \theta_2]$. Let $[\theta_1, \theta_2]$ and $[\theta_3, \theta_4]$ be two such intervals, where $\theta_3 > \theta_2$. Denote by h, h' the respective choices of overtime hours in the two pools. By the previous step, $h' \geq h$. Since the two pools are distinct, $h' < h$. Then $\pi(h) > 0$ and $\pi(h') = 1$.

First, assume by contradiction that $\pi(h) = 0$: then for $\theta \in [\theta_1, \theta_2]$, choosing the pooled overtime hours h rather than their myopic first best cannot increase their chances of promotion, and therefore cannot be optimal, which contradicts the existence of the pool in equilibrium.

Now, because of step 1, it must be that $\mathbb{E}(\theta \mid h) \geq 0$, since workers who choose h have a strictly positive probability of being promoted.

Since $\mathbb{E}(\theta \mid h_0)$ is derived using Bayes' rule for the observed values of h_0, it must be that $\mathbb{E}(\theta \mid h') > \mathbb{E}(\theta \mid h)$, since $\theta_3 > \theta_2$. Hence $\mathbb{E}(\theta \mid h') > 0$. Again by step 1, $\pi(h') = 1$.

Step 4

Any equilibrium has at least one pool.

Suppose this was not the case, and therefore that each type was picking a distinct $h(\theta)$. By step 2, this implies that the map $h(\theta)$ is strictly increasing. Because of full separation, $\mathbb{E}(\theta \mid h(\theta)) = \theta$ and therefore $\pi(h(\theta)) = 0$ for all types $\theta < 0$, and $\pi(h(\theta)) = 1$ for all types $\theta > 0$. Since all negative types are not promoted, they cannot get a higher payoff by playing an action different than their myopic first best, and therefore $h(\theta) = h^m(\theta), \forall \theta < 0$.

Case 1 If the map $h(\theta)$ is continuous as $\theta = 0$, then for a small value of ε, type $\theta = -\varepsilon$ could choose hours $h(\varepsilon)$, mimicking a higher type, and obtaining a jump in the promotion probability from $\pi(h(-\varepsilon)) = 0$ to $\pi(h(\varepsilon)) = 1$. This deviation is profitable if and only if

$$wh(-\varepsilon) - c(h(-\varepsilon), -\varepsilon) < wh(\varepsilon) - c(h(\varepsilon), -\varepsilon) + V.$$

Rearranging we get

$$\frac{c(h(\varepsilon),-\varepsilon)-c(h(-\varepsilon),-\varepsilon)}{h(\varepsilon)-h(-\varepsilon)} < w + \frac{V}{h(\varepsilon)-h(-\varepsilon)}.$$

This inequality must hold because, as ε approaches 0, the left-hand side approaches $\frac{\partial}{\partial h}c(h,-\varepsilon)$ and the right-hand side approaches infinity. Therefore, I have argued that type $-\varepsilon$ has a profitable deviation and would rather choose $h(\varepsilon)$, so we reach a contradiction.

Case 2 If the map $h(\theta)$ has a jump at $\theta = 0$, then since $h(\theta)$ is non-decreasing it must be a jump up. Since each type $\theta < 0$ is not promoted, they are are choosing their myopic first best, i.e. $h(\theta) = h^m(\theta)$ for all $\theta < 0$. Since the map $h^m(\theta)$ has no jumps, but $h(\theta)$ does, it must be that there exists a type $\bar{\theta} > 0$ such that $h(\bar{\theta}) > h^m(\bar{\theta})$. Now consider some h such that $h^m(\theta) < h < h(\bar{\theta})$, such that $h = h(\theta)$ for some θ. Since $h(\theta)$ is strictly increasing it must be that this type $\theta \in (0, \bar{\theta}]$. Since $\theta > 0$, it must be that $\pi(h) = 1$. Therefore, type $\bar{\theta}$ could increase her payoff by lowering her choice to h, since h is closer to $h^m(\bar{\theta})$ than $h(\bar{\theta})$, but maintaining the same promotion probability. This is a profitable deviation, hence a contradiction.

We have therefore argued that each equilibrium has at least one pool.

Step 5

Define a choice of overtime $h' \neq h(\theta)$ as weakly equilibrium dominated for type θ if

$$wh' - c(h',\theta) + V \leq wh(\theta) - c(h(\theta),\theta) + \pi(h(\theta))V,$$

and equilibrium dominated if the above holds strictly. In word, h' is equilibrium dominated for worker θ if the equilibrium payoff associated to $h(\theta)$ is higher than what she would get by working h' overtime hours and being promoted with certainty.

I now show that, in any equilibrium, if $h^* > h(\theta)$ is weakly equilibrium dominated for some type θ, then it is equilibrium dominated for all $\theta' < \theta$.

By definition of weak equilibrium dominance, we know that

$$wh^* - c(h^*,\theta) + V \leq wh(\theta) - c(h(\theta),\theta) + \pi(h(\theta))V,$$

Now consider $\theta' < \theta$.

$$
\begin{aligned}
wh(\theta') - c(h(\theta'), \theta') + \pi(h(\theta'))V &\geq wh(\theta) - c(h(\theta), \theta') + \pi(h(\theta))V \\
&= wh(\theta) - c(h(\theta), \theta) + \pi(h(\theta))V + c(h(\theta), \theta) - c(h(\theta), \theta') \\
&\geq wh^* - c(h^*, \theta) + V + c(h(\theta), \theta) - c(h(\theta), \theta') \\
&= wh^* - c(h^*, \theta') + V + c(h(\theta), \theta) - c(h(\theta), \theta') + c(h^*, \theta') - c(h^*, \theta) \\
&= wh^* - c(h^*, \theta') + V + [c(h(\theta), \theta) - c(h^*, \theta)] - [c(h(\theta), \theta') - c(h^*, \theta')] \\
&> wh^* - c(h^*, \theta') + V.
\end{aligned}
$$

The first step above is simply due to the fact that in equilibrium type θ' does not want to mimic type θ. The second step is just algebra (adding and subtracting $c(h(\theta), \theta)$), and the third step follows from the fact that h^* is weakly equilibrium dominated for type θ, so that type type θ would switch from $h(\theta)$ to h^* even if that means a jump in the promotion probability up to 1. Next, we add and subtract $c(h^*, \theta')$ and rearrange terms. The last step follows from the fact that $\frac{\partial^2}{\partial h \partial \theta} c(h, \theta < 0$, so the increase in the cost of working as overtime hours go from $h(\theta)$ to h^* is bigger for type $\theta' < \theta$ than it is for type θ.

Step 6

The refinement I employ, following Esteban and Ray (2006), rules out the possibility that the firm, after observing an off-path choice of overtime hours, h', forms a posterior whose support contains any type for which h' is equilibrium dominated.

I now show that every refined equilibrium can have at most one pool.

Suppose, by way of contradiction, that there are two pools, with all types in the interval $[\theta_1, \theta_2]$ choosing h^*, and all those in the interval $[\theta_3, \theta_4]$, with $\theta_2 < \theta_3$, choosing $h^{**} > h^*$. By step 3, we know that $\pi(h^{**}) > 0$. Hence, by step 1, we also know that $\mathbb{E}(\theta \mid h^*) \geq 0$. Since posterior beliefs on θ are computed using Bayes' rule whenever possible, it must be that $0 \in [\theta_1, \theta_2]$. Since type θ_2 does not want to mimic anyone, and in particular any type in the subsequent interval, it must be that

$$
\begin{aligned}
wh^* - c(h^*, \theta_2) + \pi(h^*)V &\geq wh^{**} - c(h^{**}, \theta_2) + \pi(h^{**})V \\
&= wh^{**} - c(h^{**}, \theta_2) + V
\end{aligned}
$$

where the second line follows from the fact that $\pi(h^{**}) = 1$, according to step 3. Hence,h^{**} is weakly equilibrium dominated for type θ_2, which implies, by step 5, that it is equilibrium dominated for any type $\theta \leq \theta_2$, and in particular for type 0.

Now, since all types in the interval $[\theta_3, \theta_4]$ select the same signal h^{**}, and $\theta_3 < \theta_4$, it must be that there exists one type $\theta \in [\theta_3, \theta_4]$ for which $h^m(\theta) \neq h^{**}$. Call $\bar{h}$, a choice of overtime hours that is ε-closer to $h^m(\theta)$ than h^{**} is. Since h^{**} is equilibrium dominated for type 0, we can find $\varepsilon > 0$ small enough, such that this new choice is also equilibrium dominated for type 0 i.e.

$$w\bar{h} - c(\bar{h}, 0) + V < wh^* - c(h^*, 0) + \pi(h^*)V.$$

By applying Step 5 again, $\bar{h}$ is equilibrium dominated for all $\theta' < 0$. Suppose now that type θ deviates from h^{**} to $\bar{h}$. If $\bar{h}$ is an on-path choice (i.e. it was played by some type in this equilibrium), since $\bar{h}$ is very close to $h^{**} > h^*$, and since $\mathbb{E}(\theta \mid h^{**}) > 0$ it must be, by step 1, that $\mathbb{E}(\theta \mid \bar{h}) > 0$ as well. This is because on-path posterior beliefs are formed using Bayes rule: type 0 in equilibrium is playing $h^* < \bar{h}$, and $h(\theta)$ is non-decreasing, so all types below 0 are playing actions below $\bar{h}$ as well. If instead $\bar{h}$ is an off-path choice, and it was not played by any type in equilibrium, still by the refinement concept employed the posterior belief on θ given $\bar{h}$ must exclude all types of which this is an equilibrium dominated choice, therefore it must exclude all the types $\theta' \leq 0$. Hence $\mathbb{E}(\theta \mid \bar{h}) > 0$ in this case too. It follows from step 1, that $\pi(\bar{h}) = 1$. This implies that type θ, by switching to $\bar{h}$ from h^{**}, would be promoted with the same probability, but be closer to her myopic first best, hence increasing her payoff overall.

Step 7

We have argued that every refined equilibrium has exactly one pool. Let's denote it by $[\theta_1, \theta_2]$, and let's call h^* the common action chosen by all the types in this interval. Then $\pi(h^*) = 1$.

To see this, suppose it wasn't the case and $\pi(h^*) < 1$. Then, by step 1, $\mathbb{E}(\theta \mid h^*) = 0$.

Since $\pi(h^*) < 1$, the promotion probability of any type $\varepsilon \in [0, \theta_2]$ can be increased, so that for any such ε there exists some $\bar{h} > h^*$ such that

$$wh^* - c(h^*, \varepsilon) + \pi(h^*)V = w\bar{h} - c(\bar{h}, \varepsilon) + V.$$

Hence $\bar{h}$ is weakly equilibrium dominated for type ε. By step 5, $\bar{h}$ is equilibrium dominated for all $\theta < \varepsilon$. Now, it must be that $\mathbb{E}(\theta \mid \bar{h}) > 0$: see step 6 for the discussion of why this is the case both if $\bar{h}$ is on path and if it isn't. This means, by step 1, that $\pi(\bar{h}) = 1$. Hence, any type $\theta \in (\varepsilon, \theta_2)$ who deviates to this $\bar{h} > h^*$, a choice of overtime hours that is equilibrium dominated for all $\theta < \varepsilon$, will be promoted for sure. Then, type θ has a profitable deviation from h^* to $\bar{h}$:

$$\begin{aligned} wh^* - c(h^*, \theta) + \pi(h^*)V &< w\bar{h} - c(\bar{h}, \theta) + V \\ c(\bar{h}, \theta) - c(h^*, \theta) w(\bar{h} - h^*) + (1 - \pi(h^*)V)V \\ c(\bar{h}, \theta) - c(h^*, \theta) c(\bar{h}, \varepsilon - c(h^*, \varepsilon). \end{aligned}$$

Notice that the last step above comes from the indifference of type ε, and that we know that the final inequality holds because of the negative cross-derivative of the cost function.

This means that type θ can get a higher payoff by deviating to $\bar{h}$, and being promoted for sure. Hence $\pi(h^*) = 1$.

Step 8

We have argued that every refined equilibrium has exactly one pool, $[\theta_1, \theta_2]$, with common action h^*, and that the promotion probability in such pool is equal to 1. I now characterize the unique equilibrium in this set by showing that $h^m(\theta_2) \leq h^*$, with equality if the interval $[\theta_1, \theta_2]$ is interior to the support of θ.

Suppose, by way of contradiction, that $h^m(\theta_2) \neq h^*$. Consider first the case in which $h^* < h^m(\theta_2)$. I claim that $\mathbb{E}(\theta \mid h) > 0$ for all $h > h^*$. This is because if such h is on path, it must be that $h = h^m(\theta)$, for some $\theta > \theta_2$. If h is off-path, then we apply the same argument from Step 6 to argue that $\theta = 0$ cannot be in the support of the firm's belief on θ. Since $\mathbb{E}(\theta \mid h) > 0$, then $\pi(h^{**}) = 1$. So type θ_2 has a profitable deviation to some $h > h^*$, and closer to their myopic first best $h^m(\theta_2)$, since promotion probability is still equal to 1. So we know that $h^* \geq h^m(\theta_2)$.

If $h^* > h^m(\theta_2)$, then, if θ_2 is interior, $\forall \theta > \theta_2$, we claim that $h(\theta) > h^*$: this is because by step 2 the map $h(\theta)$ is non-decreasing, and since there is no other pool (by step 6), $h(\theta)$ is actually strictly increasing to the right of θ_2. Now, consider a type $\theta > \theta_2$, but close enough so that $h^m(\theta) < h^*$ as well (such type exists since $h^m(\theta)$ is continuous). Since $\pi(h^*) = 1$, type θ cannot have higher probability of promotion than type θ_2. Hence, she could lower her overtime hours, moving her choice closer to $h^m(\theta)$, without decreasing her promotion probability.

Step 9

I now show that $\mathbb{E}(\theta \mid h^*) = 0$.

By step 1, $\mathbb{E}(\theta \mid h^*) \geq 0$, since $\pi(h^*) = 1$. If $\mathbb{E}(\theta \mid h^*) > 0$, then it means that $0 \notin [\theta_1, \theta_2]$. Since $0 > \theta_2$ would lead to $\mathbb{E}(\theta \mid h^*) < 0$, it must be that $0 < \theta_1$. Notice that $h^m(\theta_1) < h^*$:

this follows from the fact that $h^m(\theta_2) = h^*$ (step 8), and $\theta_1 < \theta_2$. Also, by step 1, we know that $\forall \theta \in (0, \theta_1)$, $\pi(h(\theta)) = 1$. Now let's lower $h(\theta_1)$ from h^* to some h' closer to $h^m(\theta_1)$. If some other type $\theta < \theta_1$ is choosing h' in equilibrium, then $\pi(h') = 1$. If not, we can apply a similar argument as in step 6 to argue that $(E)(\theta \mid h') > 0$ under the refined equilibrium concept, and therefore $\pi(h') = 1$ in this case too. Hence, type θ_1 can lower his action, and get closer to his myopic first best, without renouncing her promotion probability of 1, which contradicts the fact that $h(\theta_1) = h^*$.

This shows that $\mathbb{E}(\theta \mid h^*) = 0$.

To conclude, I have therefore showed that the unique equilibrium of the refined game can be characterize by the system of equations defined in Proposition 1.

Notice the assumption that w is low enough so that the firm always wants to promote workers with $\mathbb{E}(\theta) \geq 0$, ruling out the instance where achieving this separation is "too costly" in terms of overtime payments to the workers. In other words, my equilibrium definition only imposes that the firm is maximizing the stage-2 profits. Given the assumption that $w < 1$, the firm wants to maximize overtime hours in the first period, beside wanting to select the best workers for promotion: however, absent commitment the two motives are not in contrast. This is because the firm will never promote, at the beginning of stage 2, a worker who picked h such that $\mathbb{E}(\theta \mid h) < 0$ (this follows from the absence of commitment), and in the equilibrium derived I have already proved that $\mathbb{E}(\theta \mid h^*) = 0$, so that including any other type in the pool would lower the promotion probability of the h^*, violating the equilibrium conditions.

B Proof of Proposition 2

We want to show that as w decreases, the quality of promoted workers increases, i.e. $\mathbb{E}(\theta \mid \text{promoted}) = \mathbb{E}(\theta \mid \theta \geq \theta_1)$ goes up, i.e. that θ_1 goes up.

The proof is going to proceed in steps. First, I argue that as w increases θ_1, defined as the lowest type who is willing to mimic a fixed θ_2, decreases. To study this point, we need to compare two effects: first, as w increases, mimicking a fixed h^* is now cheaper, which leads θ_1 to decrease; second, the action required to mimic is now higher (because $h^* = h^m(\theta_2)$ increases with w), and this leads θ_1 to increase. Appropriate conditions on the cost function $c(h, \theta)$ guarantee that the second effect does not dominate the first one, and that overall w and θ_1 are inversely related. Second, I show that $\mathbb{E}(\theta \mid h^m(\theta_2)$ is increasing in θ_2. Finally, I show that $\mathbb{E}(\theta \mid h^m(\theta_2)$ is inversely related to w, for a fixed θ_2, meaning that as w goes down the set of type that chooses a fixed $h^m(\theta_2)$ is better. This implies that the value of h^* such that $\mathbb{E}(\theta \mid h^*) = 0$ will be lower. Since θ_2 is defined as the type such that $h^m(\theta_2) = h^*$, this also implies that θ_2 is lower. The final step of the proof argues that this implies that θ_1 must increase.

Step 1

To study how θ_1 changes with w, remember that it is defined by the following expression:

$$wh^m(\theta_1) - c(h^m(\theta_1), \theta_1) = wh^m(\theta_2) - c(h^m(\theta_2), \theta_1) + V,$$

which simply means she is indifferent between pooling at $h^m(\theta_2)$ (and being promoted for sure) and choosing her myopic first best (and not be promoted).

The total derivative of this expression with respect to w is given by:

$$\frac{\partial \theta_1(h^m(\theta_2), w)}{\partial w} = \underbrace{\frac{\partial \theta_1(h^m(\theta_2), w)}{\partial h^m(\theta_2)}}_{\text{Term 1}} \underbrace{\frac{\partial h^m(\theta_2)}{\partial w}}_{\text{Term 2}} + \underbrace{\frac{\partial \theta_1(h^m(\theta_2), w)}{\partial w}}_{\text{Term 3}}.$$

Deriving them one by one we notice that the effect of a change in the action to mimic is given by

$$\frac{\partial \theta_1(h^m(\theta_2), w)}{\partial h^m(\theta_2)} = -w + \frac{\partial}{\partial h^m(\theta_2} c(h^m(\theta_2), \theta_1).$$

As for Term 2, i.e. $\frac{\partial}{\partial w} h^m(\theta_2)$, remember that

$$w = \frac{\partial}{\partial h} c(h^m(\theta_2), \theta_2),$$

so by differentiating on both sides we get

$$\begin{aligned} 1 &= \frac{\partial}{\partial w} \frac{\partial}{\partial h} c(h^m(\theta_2), \theta_2) \\ &= \frac{\partial}{\partial h} \frac{\partial}{\partial h} c(h^m(\theta_2), \theta_2) \frac{\partial}{\partial w} h^m(\theta_2) \\ \frac{\partial}{\partial w} h^m(\theta_2) &= \frac{1}{\frac{\partial^2}{\partial^2 h} c(h^m(\theta_2), \theta_2)}. \end{aligned}$$

Finally, term 3 is the effect of a change in w holding fixed the action to imitate, so

$$\frac{\partial \theta_1(h^m(\theta_2), w)}{\partial w} = h^m(\theta_1) + \frac{\partial}{\partial w} h^m(\theta_1) w - \frac{\partial}{\partial w} h^m(\theta_1) \underbrace{\frac{\partial}{\partial h^m(\theta_1)} c(h^m(\theta_1), \theta_1)}_{=w} - h^m(\theta_2) = h^m(\theta_1) - h^m(\theta_2)$$

Plugging terms back in we get

$$\frac{\partial\theta_1(h^m(\theta_2),w)}{\partial w} = \left(-w + \frac{\partial}{\partial h^m(\theta_2}c(h^m(\theta_2,\theta_1)\right)\frac{1}{\frac{\partial^2}{\partial^2 h}c(h^m(\theta_2),\theta_2)} + h^m(\theta_1) - h^m(\theta_2).$$

This expression is negative if and only if

$$h^m(\theta_1) - h^m(\theta_2) < \left(w - \frac{\partial}{\partial h^m(\theta_2)}c(h^m(\theta_2,\theta_1)\right)\frac{1}{\frac{\partial^2}{\partial^2 h}c(h^m(\theta_2,\theta_2)}.$$

Plugging in the definition of w as $\frac{\partial}{\partial h}c(h^m(\theta_1),\theta_1)$ and moving terms around the condition becomes

$$\frac{\partial^2}{\partial^2 h}c(h^m(\theta_2),\theta_2) > \frac{\frac{\partial}{\partial h}c(h^m(\theta_1),\theta_1) - \frac{\partial}{\partial h}c(h^m(\theta_2),\theta_1)}{h^m(\theta_1) - h^m(\theta_2)}.$$

By the mean value theorem, the right hand side is equal to

$$\frac{\partial}{\partial h}\frac{\partial}{\partial h}c(\tilde{h},\theta_1)$$

for some $\tilde{h} \in [h^m(\theta_1), h^m(\theta_2)]$, so that the overall condition becomes

$$\frac{\partial^2}{\partial^2 h}c(h^m(\theta_2),\theta_2) > \frac{\partial^2}{\partial^2 h}c(\tilde{h},\theta_1).$$

If $\frac{\partial^3}{\partial^3 h}c(h,\theta) > 0$, then $\frac{\partial^2}{\partial^2 h}c(h^m(\theta_2,\theta_1) > \frac{\partial^2}{\partial^2 h}c(\tilde{h},\theta_1)$, simply because $h^m(\theta_2) > \tilde{h}$ and the second derivative is increasing. Hence the expression we are after is implied by

$$\frac{\partial^2}{\partial^2 h}c(h^m(\theta_2),\theta_2) \geq \frac{\partial^2}{\partial^2 h}c(h^m(\theta_2),\theta_1) > \frac{\partial^2}{\partial^2 h}c(\tilde{h},\theta_1).$$

If the second derivative of $c(h,\theta)$ is (weakly) increasing in θ, then the above expression follows.

Hence, sufficient condition for the direct effect to dominate are given by:

$$\frac{\partial^3}{\partial^3 h}c(h,\theta) > 0$$

and, at the same time

$$\frac{\partial}{\partial\theta}\frac{\partial^2}{\partial^2 h}c(h,\theta) \geq 0$$

If these two technical condition jointly hold, then as w increases the initial type θ_1 that was indifferent about imitating a fixed type θ_2 now strictly prefers to imitate θ_2 (even if

$h^m(\theta_2)$ is now higher), so that the marginal type that is willing to mimic θ_2 goes down.

An example of a class of cost function that satisfies these two conditions, as well as the standard assumptions made in section 2.1, is given by $c(h, \theta) = f(h) + g(\theta)h$, for $f(h)$ positive, increasing, convex and increasing convex, and $g(\theta)$ positive and decreasing.

Step 2

The expected value of θ over the pooling interval increases with θ_2.

As θ_2 increases, so does $h^* = h^m(\theta_2)$, by definition of myopic optimum. Hence, θ_1 increases (the action to mimic has gone up, so the marginal guy who is willing to do it is better selected), and therefore $\mathbb{E}(\theta \mid h^*) = \mathbb{E}(\theta \mid \theta \in [\theta_1, \theta_2]$ increases as well.

Step 3

For a fixed θ_2, $\mathbb{E}(\theta \mid h^m(\theta_2)$ increases with w. This is because for a higher w, the marginal type that is willing to mimic θ_2 is lower (by step 1 of this proof): if the lower bound of the expectation decreases, the expected value decreases as well. Hence, the value of θ_2 such that $\mathbb{E}(\theta \mid h^m(\theta_2) = 0$ must decrease if w decreases.

Step 4

In the initial equilibrium $\mathbb{E}(\theta \mid [\theta_1, \theta_2] = 0$. As w increases, θ_2 increases, by steps 1-3. Since in the new equilibrium $\mathbb{E}(\theta \mid [\theta_1', \theta_2'] = 0$, it must be that θ_1 increases.

C Discussion of the boundary case

The purpose of this Appendix is to discuss how Proposition 1 and 2 are modified in the case in which, the equilibrium entails $\theta_2 = \bar{\theta}$. For Proposition 1 the only thing to notice is that the property $h^m(\theta_2) = h^*$ does not hold anymore, and in particular $h^m(\theta_2) < h^*$. The quantity h^*, i.e. the pooling action, needs to be above $h^m(\theta_2)$, and it is still uniquely determined so that θ_1, the lowest types that is willing to pool at h^*, is high enough to generate $\mathbb{E}(\theta \mid h^*) = 0$.

As for Proposition 2, it goes true in the exact same way if we are at the boundary case, except for the observation that if h^* is not anymore defined as $h^m(\theta_2)$ is still does change so that the marginal type who is willing to choose h^* is such that $\mathbb{E}(\theta \mid h^*) = 0$.

D Details on the cleaning procedure

This Appendix describes in detail how I constructed by sample, and it includes some descriptive statistics and balance tables.

Starting from the universe of Portuguese individuals who work in private sector firms with at least one wage earner, I restrict the sample in the following way. Using the waves from 2006 to 2014, I first drop individuals whose employment status is not labeled as "employee" (hence, I drop employers, unpaid family workers, members of coops, and people who can't be classified), and individuals who in a given month where not paid their usual remuneration, meaning people for which the recorded normal earnings are less than what they would get in a normal month (for example, because they were on leave for part of the month). Finally, I keep only workers in firms whose one-digit industry code is either retail or hospitality. Next, I drop workers who are not covered by any union contract, which is less than 3% in retail and hospitality.[44] For these workers, overtime premium is not determined by collective bargaining, and it is therefore unclear how their overtime pay would be affected by the reform.

My analysis focuses on entry-level workers: I select them based on the rung of the hierarchy ladder in which they are in 2011. This implies removing from the sample all the workers that are not in the data in 2011, and those who are in the data but not in one of the three bottom rungs.

Table D.1 shows all the cleaning steps, together with the share of the sample lost in each of them.

Table D.2, D.3 and D.4 compare the samples before and after each of the three last cleaning steps. Table D.2 compares the workers who, in 2011, were in either of the three bottom rungs of the firm latter, with those who were in any other rung. Table D.3 shows how the sample for which the treatment variable is defined differs from the rest. The treatment variable is not defined if in a given collective contract and a a given occupation, nobody does overtime in the any of the years 2009-2011. Table D.4 shows the impact of the treatment variable cleaning: it compares the observations in my sample with the ones that are dropped because of anomalies in the treatment variable. The anomalies considered are of two different types. First, exclude from my analysis observations for which the treatment variable is below 50, i.e. the average overtime premium in a given contract-occupation, averaged across the years 2009-2011, was below the Labor Code default of 50%, which according to Martins (2017) acted as a minimum. Second, I exclude from my analysis observations for which the contract-occupation average decreased going from the

[44] This figure is around 10% in the whole private sector.

2009-2011 average to the 2012-2014 average.

Finally, Table D.5 describes the final sample according to all the main variables used in the analysis.

Table D.1: Cleaning steps

Sample description	# Workers	Share dropped
Universe 2006-2014	4,720,568	.
Only "employees"	4,498,002	4.7 %
Only full remuneration	4,089,764	9.1 %
Only retail and hospitality	1,352,344	66.9 %
Only covered by union contract	1,322,486	2.2%
Present in 2011	589,991	55.4%
Only entry-level in 2011	428,844	27.3%
Only with non-missing treatment variable	421,196	1.8%
Only with consistent treatment variable	406,222	3.6%

Table D.2: Balance table: Entry level workers

	Not entry-level			Entry-level				
	N	Mean	SD	N	Mean	SD	Diff.	SE
Female	161147	0.45	(0.50)	428844	0.52	(0.50)	0.06***	(0.00)
Tenure	161140	8.56	(8.32)	428835	7.44	(7.46)	-1.12***	(0.01)
Normal hourly wage	161147	7.01	(6.47)	428844	3.72	(1.54)	-3.30***	(0.00)
Overtime premium	23845	81.56	(45.07)	76037	85.15	(47.46)	3.59***	(0.23)
Overtime hours	161147	0.94	(5.56)	428844	0.79	(4.55)	-0.14***	(0.01)
Avg annual wage growth	161147	3.13	(5.82)	428844	2.41	(3.69)	-0.72***	(0.01)
	161147	0.06	(0.23)	428844	0.05	(0.22)	-0.00***	(0.00)
>3 employees	161147	0.85	(0.36)	428844	0.82	(0.39)	-0.03***	(0.00)
Treatment defined	161147	0.96	(0.19)	428844	0.98	(0.13)	0.02***	(0.00)

Note: This table compares individual who are entry-level workers in 2011 (and therefore are retained in my sample) with those who are not. Entry level workers are defined as being either non-skilled, or skilled, or semi-skilled.

Table D.3: Balance table: treatment variable cleaning

	Treat not defined			Treat defined				
	N	Mean	SD	N	Mean	SD	Diff.	SE
Female	7614	0.33	(0.47)	421230	0.52	(0.50)	0.19***	(0.00)
Tenure	7613	8.72	(8.47)	421215	6.74	(7.24)	-1.97***	(0.04)
Age (coded 17-68)	7605	40.40	(10.82)	421132	37.05	(10.61)	-3.35***	(0.06)
Normal hourly wage	7614	3.97	(1.84)	421230	3.63	(1.50)	-0.34***	(0.01)
Regular bonuses	7614	92.42	(118.65)	421230	98.16	(157.92)	5.74***	(0.95)
Avg annual wage growth	7614	2.24	(4.15)	421230	2.46	(3.72)	0.22***	(0.02)
Promoted	7614	0.04	(0.19)	421230	0.06	(0.24)	0.03***	(0.00)
>3 employees	7614	0.72	(0.45)	421230	0.72	(0.45)	-0.00	(0.00)
Irregular bonuses	7614	49.82	(220.81)	421230	62.54	(1417.87)	12.72	(8.45)
Treatment defined	7614	0.00	(0.00)	421230	1.00	(0.00)	1.00	(0.00)

Note: This table compares individual who have a non-missing value of the treatment variable (and therefore are retained in my sample) with those who do not. The treatment variable is missing if none of the workers in a given contract-occupation does overtime in any of the years from 2009 to 2011.

Table D.4: Balance table: consistent treatment variable

	In sample			Not in sample				
	N	Mean	SD	N	Mean	SD	Diff.	SE
Female	398628	0.53	(0.50)	22602	0.40	(0.49)	-0.12***	(0.00)
Tenure	398619	7.16	(7.36)	22602	7.43	(7.36)	0.28***	(0.02)
Normal hourly wage	398628	3.68	(1.52)	22602	3.75	(1.60)	0.07***	(0.00)
Overtime premium	69438	89.02	(47.27)	2471	57.03	(40.77)	-31.99***	(0.66)
Overtime hours	398628	0.84	(4.64)	22602	0.67	(4.42)	-0.17***	(0.01)
Avg annual wage growth	398628	2.42	(3.71)	22602	2.40	(3.70)	-0.03**	(0.01)
Promoted	398628	0.07	(0.26)	22602	0.05	(0.22)	-0.02***	(0.00)
>3 employees	398628	0.73	(0.44)	22602	0.68	(0.47)	-0.05***	(0.00)
Treatment defined	398628	1.00	(0.00)	22602	1.00	(0.00)	0.00	(0.00)

Note: This table compares individual who have a value of the treatment variable that satisfies minimal legal consistency requirement (and therefore are retained in my sample) with those who do not. In particular, the inconsistencies that cause an exclusion for the sample are either a value of the treatment variable below 50, or a decrease in the average overtime premium in the contract-by-occupation around 2012.

Table D.5: Summary statistics

	N	Mean	SD	Min	Max	IDs
Female	2392105	0.52	0.50	0	1	406242
Tenure	2391593	7.43	7.46	0	62	406233
Normal hourly wage	2392105	3.71	1.54	1.78	127.9	406242
Overtime premium	158723	86.2	47.4	-2.56	300.0	73464
Overtime hours	2392105	0.80	4.55	0	176	406242
Avg annual wage growth	2392105	2.41	3.69	-40.7	413.9	406242
Promoted	2392105	0.070	0.25	0	1	406242
>3 employees	2392105	0.72	0.45	0	1	406242
Treatment defined	2392105	0.98	0.13	0	1	

Note: This table reports summary statistics for all the variables used in the analysis.

E Effects of the reform on various outcomes

This Section shows the results of estimation equation (3) on different outcomes.

In particular, Table E.7 confirm that treated workers were not compensated for the decrease in the overtime wage by any increase in the normal hourly wage.

Table E.6: Effects of overtime reform: secondary outcomes

	(1) Promoted	(2) Same firm	(3) Total earnings
$\text{Post}_t \times \text{Treatment Intensity}_j(i)$	0.0300	0.0359	-5.788
	(0.0402)	(0.0410)	(5.897)
Observations	2,348,005	2,348,005	2,348,005
Unique IDs	398,546	398,546	398,546
Year FEs	Yes	Yes	Yes
Contract-by-Rung FEs	Yes	Yes	Yes
Mean dependent variable ($\text{Post}_t = 0$)	0.0240	0.639	721.1

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the effect of the 2012 reform on some secondary outcomes, estimated according to Equation (3). The dummy "Same firm" is equal to 1 if the individual is working in the same firm for two consecutive years.

Table E.7: Effects of overtime reform: secondary outcomes

	(1) Normal hourly wage	(2) Regular bonus	(3) Irregular bonus
$\text{Post}_t \times \text{Treatment Intensity}_{j(i)}$	-9.18e-18	3.644	-12.72
	(6.35e-17)	(2.874)	(11.69)
Observations	2,348,005	2,348,005	2,348,005
Unique IDs	398,546	398,546	398,546
Year FEs	Yes	Yes	Yes
Contract-by-Rung FEs	Yes	Yes	Yes
Mean dependent variable ($\text{Post}_t = 0$)	3.636	98.25	62.23

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the effect of the 2012 reform on some secondary outcomes, estimated according to Equation (3).

Table E.8: Effects of overtime reform on firms' performance

	(1) Hiring rate	(2) Separation rate	(3) Sales
$\text{Post}_t \times \text{Treatment Intensity}_i$	0.0000504	-0.0000407	-123.8
	(0.0000292)	(0.0000288)	(604.8)
Observations	525,034	525,034	524,488
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Mean dependent variable	0.0345	0.0524	137712.0

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the effect of the 2012 reform on firm-level outcomes. These are estimated on the firm-equivalent version of Equation (3): treatment intensity of the firm is defined as the average treatment intensity of its workers. Controls include share of women and average normal hourly wage in the firm. Standard errors are clustered at the firm level.

Table E.9: Supporting evidence: firm level outcomes

	(1) Overtime premium	(2) Overtime hours (total)	(3) Overtime hours (average)	(4) Gap overtime hours	(5) Promotion Probability	(6) Promotion Probability if Overtime Hours == 0
$\text{Post}_t \times \text{Treatment Intensity}_i$	-0.626***	-0.0638*	-0.00221***	-0.00463***	-0.000133**	-0.000144**
	(0.0638)	(0.0269)	(0.000589)	(0.00127)	(0.0000430)	(0.0000514)
Observations	18,494	525,034	525,034	525,034	525,034	401,093
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dependent variable	83.30	4.015	0.443	0.704	0.0493	0.0749

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows the effect of the 2012 reform on firm-level outcomes. These are estimated on the firm-equivalent version of Equation (3): treatment intensity of the firm is defined as the average treatment intensity of its workers. Controls include share of women and average normal hourly wage in the firm. Standard errors are clustered at the firm level.