Essays in Industrial Organization

Ph.D. Thesis

Charles Pébereau

Toulouse School of Economics

30th August 2021

Acknowledgements

The story starts with a continental breakfast in Paris and ends with a one-way ticket to California. Most of it happened in Toulouse, but not only. There was a (planned) trip in New Zealand, three (unplanned) confinements in Tarn, Gers and Bretagne, a new university building (planned but late), and multiple (fortuitous) work trips in Portugal in the summer. This is not the synopsis of a spy movie but an overview of my PhD in economics. And to prove that I am not a secret agent I will now publicly thank my contacts.

My first thoughts go to Thomas-Olivier Léautier, whom I first met at a breakfast after hearing of him at a dinner. I was contemplating the idea of pursuing a PhD to acquire scientific expertise but I was unfamiliar with Economics - so I asked for his advice. I suggested a quick PhD, with an industrial partner, in Paris. He invited me to take a second croissant ... and the journey started. We have only briefly worked together but crossing your path changed the course of my life and I am very grateful.

I wish to express my sincere gratitude to my advisor and mentor, Jean Tirole, for the greatest gift of all: a breathtaking intellectual adventure! The journey was all the more exciting that the mission order was unexpected - and unusual. In a nutshell, he recommended me to seek inspiration in science fiction to identify some of our greatest fears and address them by mixing seminal concepts and state-of-the-art methodology. I am thankful and honored for this great challenge, and for your guidance and watchful eye at every step. Working aside you has been inspiring and your curiosity and enthusiasm for novelty are contagious. I am now going to explore the Silicon Valley - the 'tech capital of the world' - but I am taking some white chalk and a few sci-fi books with me.

This PhD has also been a fantastic experience thanks to the support and the collaboration with a large and friendly research community at TSE and elsewhere. My project on real-time electricity pricing - now a chapter of this thesis - is an emblematic example. Thomas-Olivier Léautier put me in contact with the Electricity Authority in New Zealand, Natalia Fabra encouraged me to work with them, Estelle Cantillon flew there with me, Phil Bishop and his team welcomed me, Stefan Ambec helped me find funding and helped me all along the project, Mathias Reynaert provided irreplaceable guidance, and Kevin Remmy worked aside me as a co-author! I am also particularly grateful to Johannes Horner and Yassine Lefouilli for advising me throughout the PhD on all my projects and helping me sharpen my skills and critical abilities, and to Jacques Crémer and Bruno Jullien for regularly giving me feedback as well as career advice. That is just the tip of the iceberg because there are also all the researchers whom I have not met but who inspired me. In that respect, I wish to thank Heski Bar-Isaac and Mar Reguant for accepting to be members of my committee, I look forward to the discussion.

These past six years represent, for me, the discovery of a discipline, Economics, of a city, Toulouse, but also of an incredible diversity of people. To my greatest pleasure, somehow, between the flatshares, the coffee breaks, the confinements and the trips, work life and social life got tangled up. Friends and colleagues, you have been an inexhaustible source of inspiration and joy. I have a special thought for Sarah, a dear friend and travel companion - spiritual, intellectual, and worldwide - and for Jeremy, Peter and Filip with whom we have formed an unexpected quartet. Finally, I want to thank my family for their precious love and their warm support.

Summary - English

This thesis consists of three independent chapters borrowing concepts and methods from the field of industrial organization to address some of the challenges that have arisen with the development of information technologies.

The first chapter is motivated by the growing demand for better control over who can access personal data and for which purposes - in particular concerning the spreading practice, both by public authorities and private companies, of scoring individuals. While disclosing information makes individuals accountable for their actions and can protect others from undesirable encounters, it also increases ostracism and costly avoidance strategies. I study the effecient disclosure of information by a principal about an agent who has reputation concerns vis-à-vis multiple audiences with heterogeneous valuations for information. Disclosing information to an audience jeopardizes the agent's reputation and thus incentivizes. To maximize effort under a participation constraint it is optimal to disclose undifferentiated information across the audiences who value information the most and to protect the agent's reputation vis-à-vis the other ones. When the principal internalizes the audiences' payoffs, he trades off providing more information to the most interested audiences with protecting the agent's reputation vis-à-vis the other ones. The results can rationalize existing laws protecting workers and consumers and inform the debate regarding social scoring or the sharing of information between firms and within multi-sided platforms.

In the second chapter, I study how a platform can exploit personal information to decide who to accept and who meets whom and encourage pro-social behavior. I build a model with the assumption that some individuals are better match partners than others - for instance, because they behave more altruistically - and ask who to match them with. There are no monetary transfers to provide incentives but, instead, the platform can use the best match partners as a reward for good ratings. I characterize the matching mechanism maximizing welfare on the platform. This mechanism is not entry-proof, though, and competition between platforms leads to stratification whereby the best match partners are matched together, a phenomenon

called positive assortative matching (PAM) in the literature. Thus, platform entry creates a hold-up problem which reduces individuals' provision of effort. This issue is exacerbated with data portability because it allows an entrant to screen individuals more efficiently. The framework developed is relevant for the regulation of algorithms developed by platforms of the gig economy and can be adapted to study the effect of urban planning on community harmony.

In the third chapter, co-authored with Kevin Remmy, we study empirically the introduction of real-time pricing (RTP) in New Zealand - a new electricity tariff relying on smart meters which record detailed consumption data - and our goal is to identify barriers to its adoption. While economic theory predicts that introducing RTP in an economy with rational and perfectly informed agents will lead the retail market to unravel, less than 1.25% of residential consumers adopted this tariff more than seven years after its introduction. Under this tariff, consumers are exposed to half-hourly varying spot prices, which are uncertain and volatile. We show that when ongoing spot prices spike, prospective adopters forego adoption and recent adopters switch to another tariff or reduce their electricity consumption - which is a sign of present bias. However, with experience, consumers on real-time pricing gradually become less sensitive to ongoing spot prices. We propose remedies to foster widespread adoption of real-time pricing.

Résumé - Français

Cette thèse, constituée de trois chapitres, mobilise des outils d'économie industrielle pour aborder certains des défis apparus avec l'essor des technologies de l'information.

Le premier chapitre est motivé par la demande croissante d'un meilleur contrôle concernant l'accès à et l'utilisation des données personnelles. Si transparence rend les individus responsables de leurs actions et peut protéger de rencontres indésirables, elle augmente également l'ostracisme et les stratégies d'évitement coûteuses. J'étudie, pour divers objectifs, quelle est la stratégie efficace de divulgation d'informations par un principal à propos d'un agent qui se préoccupe de sa réputation vis-à-vis de multiples publics. Divulguer de l'information met en péril la réputation de l'agent et l'incite donc à l'effort, et l'effort exercé augmente avec la quantité d'information révélée. Pour maximiser l'effort, sous une contrainte de participation, il est optimal de divulguer des informations indifférenciées à un ensemble de publics pour qui ces informations sont le plus utiles et de protéger la réputation de l'agent vis-à-vis des autres publics. Lorsque le principal internalise le bien-être de tous les publics, la divulgation n'est plus indifférenciée. Au contraire, il fournit d'autant plus d'informations à un public qu'elles lui sont utiles et, en contrepartie, protège mieux la réputation de l'agent vis-à-vis des autres publics. Les résultats peuvent rationaliser les lois existantes pour la protection des travailleurs et des consommateurs, et alimenter le débat sur la notation sociale ou encore le partage d'informations entre entreprises ou au sein des plateformes multifaces.

Dans le deuxième chapitre, j'étudie de quelle manière une plateforme peut exploiter les informations personnelles pour encourager des comportements pro-sociaux. Je construis un modèle en partant de l'hypothèse que certains individus sont des partenaires plus désirables que d'autres - par exemple parce qu'ils ont un comportement plus altruiste. Il n'y a pas de transferts monétaires pour fournir des incitations mais la plateforme peut utiliser l'appariement avec les partenaires les plus désirables comme récompense. Je caractérise le mécanisme d'appariement qui maximise le bien-être sur la plateforme. Ce mécanisme n'est cependant pas robuste contre l'entrée d'un concurrent et la concurrence pousse les plate-

formes à apparier les individus les plus désirables ensemble. L'entrée d'une nouvelle plateforme crée donc un problème de hold-up qui réduit l'effort fourni par les individus. Ce problème est exacerbé par la portabilité des données, car elle permet à l'entrant de sélectionner les individus plus efficacement. Le modèle est pertinent pour réfléchir à la régulation des algorithmes développés par les plateformes de l'économie à la tâche et peut être adapté pour étudier l'effet de la planification urbaine sur l'harmonie communautaire.

Dans le troisième chapitre, co-écrit avec Kevin Remmy, nous étudions l'introduction de la tarification en temps réel de l'électricité en Nouvelle-Zélande - un nouveau tarif reposant sur des compteurs intelligents qui enregistrent des données de consommation détaillées – afin d'identifier les obstacles à son adoption. Alors que la théorie économique prévoit que cette tarification sera rapidement adoptée par une large proportion des ménages, en Nouvelle-Zélande moins de 1,25 % des consommateurs résidentiels l'ont adopté plus de sept ans après son introduction. Avec ce tarif, les consommateurs sont exposés aux prix du marché qui varient toutes les demi-heures, et qui sont incertains et volatils. Nous montrons que lorsque les prix s'envolent, l'adoption baisse et l'attrition augmente, notamment parmi les consommateurs ayant récemment adopté. Cependant, avec de l'expérience, les consommateurs deviennent progressivement moins sensibles aux variations de prix de court terme. Nous proposons des remèdes pour favoriser l'adoption de la tarification en temps réel.

Table of Contents

Chapter 1

None of your business! Efficient disclosure policies with heterogeneous audiences

Abstract

While disclosing information makes individuals accountable for their actions and reduces adverse selection, many laws restrict access to personal data. This paper studies the efficient disclosure of information by a principal about an agent who has reputation concerns vis-à-vis multiple audiences with heterogeneous valuations for information. Disclosing information to an audience jeopardizes the agent's reputation and thus incentivizes effort and disclosing more information increases effort. To maximize effort under a participation constraint, it is optimal to disclose undifferentiated information across a subset of audiences who value information the most and to protect the agent's reputation vis-à-vis the other ones. When the principal internalizes the audiences' payoffs, he tailors the policy and discloses proportionally more information to audiences who value it more. The results can rationalize existing laws protecting workers and consumers and inform the debate regarding social scoring or the sharing of information between firms and within multi-sided platforms.

1.1 Introduction

Should our employer know as much about us as our friends? What about our banker and our landlord? Can the argument that some information is 'none of your business' be defended on efficiency grounds? With the advances in information technologies the collection and sharing of personal data, both online and offline, has become pervasive and there is a growing demand for more control over who can access what personal information and for which purposes. In a proposal for AI regulation, the European Commission calls for banning social scores by public authorities because of the risks of ostracism or discrimination in contexts which are unrelated to the context in which the information was collected. Meanwhile, several laws and social norms already address this issue by restricting access or even by differentiating the information disclosed to each audience. For instance, in many countries the law prevents some the firms, but not all, from asking or using certain personal information about their employees or job candidates - such as financial credit reports[1], criminal records[2], social media accounts[3], or drug tests[4] - with exceptions for jobs with high risks of corruption or those involving vulnerable populations[5]. Also, in France, sharing blacklists of fraudulent customers is only allowed between the firms within the same sector of activity but not across sectors[6]. Finally, several social norms dictate what is admissible gossip or how to introduce an acquaintance to different groups, such as friends or colleagues. The purpose of this paper is to investigate whether such laws and norms can be rationalized as the outcomes of an information design problem.

The literature on incomplete information has shown how disclosing information to an audience can both improve market efficiency under adverse selection and discipline agents by holding them accountable for their actions. Criminal records, blacklists, credit scores, and ratings on platforms are concrete illustrations where this dual role of information is exploited. The standard trade-off is between disciplining agents, offering them second chances, and protecting other people from undesirable encounters. The latter, reputational risk, has been studied in the context of a single audience. But the audiences - individuals and organizations - an agent will interact with in the future value access to her personal information differently.

---

[1]See Bartik and Nelson (2019), Corbae and Glover (2018), Cortes et al. (2020)

[2]See Agan and Starr (2018), Finlay (2008)

[3]See list of laws in the USA about access to social media usernames and passwords https://www.ncsl.org/research/telecommunications-and-information-technology/The firm-access-to-social-media-passwords-2013.aspx

[4]See Wozniak (2015)

[5]In France, there exist three versions for each criminal record - listing all or only some of the most severe crimes - and the firms' access to each version is regulated on the basis of how important knowing about criminal behavior is to them.

[6]A telecom service company can share a blacklist of customers who do not pay their bills with other telecom service companies but not with gas suppliers.

Yet, little is known theoretically about how to efficiently disclose information in situations with heterogeneous audiences. Accounting for this heterogeneity is the focus, and the contribution, of the paper.

Model. We build a two-period model with reputation concerns à la Holmström (1999) that we enrich with information design and heterogeneous audiences. While the model applies to different situations the lead example, throughout the paper, is the labor market. An agent may be hired by a firm in period 1, and by one of the employers $\rho \in [0, 1]$ in period 2. There is ex ante uncertainty regarding which job opportunity may arise and the identity ρ of the employer is only revealed at the beginning of the second period. There are no transfers and therefore no contracts with financial incentives. However, the firm can leverage the agent's career concerns to provide reputational incentives by strategically disclosing information to the employer.

The agent has job-specific talents, correlated between each other. Specifically, given the agent's talent for the date-1 job, $\theta \in \mathbb{R}$, her expected talent for job ρ is $\rho\theta$. Because of this correlation, information about the agent's date-1 output is useful to estimate her talent in job ρ and even more so for higher ρ. Thus, by strategically disclosing information, the firm can affect employers' hiring decisions, which incentivizes the agent to exert effort in the first period in order to avoid unemployment in the second period.

At the beginning of the game, the firm publicly commits to a disclosure policy. There are no restrictions and therefore the disclosure can be contingent on the agent's date-1 output as well as on the date-2 job opportunity. The goal is to establish how to efficiently disclose information in this environment. In particular we address the following questions. Vis-à-vis which employers should the agent' reputation be jeopardized and protected in order to maximize effort? What is the mechanism behind reputational incentives and is it affected by the existence of heterogeneous employers? How does internalizing employers' profits affect the structure of the disclosure policy? We now provide the main take-aways of the paper.

Take-aways. Efficient disclosure policies have a simple cutoff structure and the agent is recommended to an employer if and only if her date-1 output exceeds the employer-specific cutoff. In equilibrium, increasing any cutoff increases the firm's profit by incentivizing effort, but also the employers' profits by improving screening. However, it hurts the agent because effort is costly and because she is less likely to be hired. Any disclosure policy trades-off these three criteria.

In practice this cutoff structure can take several forms such as which crimes are listed on a criminal record and for how long, the number of unpaid bills it takes to be blacklisted by a telecom company, or rules for bankruptcy flags removal.

We say that the agent's reputation is protected vis-à-vis employer ρ when the disclosure policy maximizes her chances to be hired, and thus employer ρ makes zero profit. Otherwise, we say that her reputation is jeopardized vis-à-vis employer ρ. Because disclosure policies have a cutoff structure we cannot use classic notions of informativeness - such as Blackwell ranking - to measure how much information is disclosed to each employer. Instead, we use employer expected profit to proxy for informativeness. We say that a disclosure policy ψ provides more information to employer ρ than ψ' if its expected profit is strictly higher under ψ. Relatedly, consider two employers ρ and ρ' vis-à-vis which the agent's reputation is jeopardized. We say that employer ρ is more informed than employer ρ' if their expected profits are both strictly higher with the information disclosed to employer ρ.

To incentivize effort the agent's reputation is jeopardized in priority vis-à-vis employers hiring for jobs similar to the date-1 job and the cutoff is uniform across these employers. The agent's reputation is protected vis-à-vis the other employers. A higher uniform cutoff jeopardizes the agent's reputation vis-à-vis more employers and discloses more information to them. These two channels incentivize higher effort but yield lower payoff to the agent. and therefore the uniform cutoff maximizing effort if the highest one satisfying the agent's participation constraint.

Employers who hire for jobs similar to the date-1 job value output information more because it is more useful to assess the agent's talent for these jobs. Therefore, a uniform cutoff is no longer optimal when employers profits are internalized. Among employers vis-à-vis which the agent's reputation is jeopardized, those who value information the most receive more information; that is, the cutoff increases with ρ. Such differentiated disclosure dampens incentives but improves screening.

The paper proceeds as follows. Section 1.2 reviews the literature. We present the model in Section 1.3, solve it in the case with a single audience in Section 1.4 and in the case with heterogeneous audiences in Section 1.5. In Section 1.7 we propose extensions and discuss the results with cases studies.

1.2 Related literature

Compared with the existing literature, our innovation is to combine information design and audience heterogeneity in an environment with reputational concerns. Our focus is on contractual efficiency and we emphasize the limitations stemming from the existence of multiple audiences to which information can be disclosed but who value information differently.

Implicit and explicit incentives. First and foremost, our paper is a synthesis between the explicit and implicit incentive models of Hölmstrom (1979) and Holmström (1999) that we complement with techniques borrowed from the information design literature - see below - and allow for heterogeneous audiences. We compute the equilibrium contracts between a principal and an agent in the presence of moral hazard as in Hölmstrom (1979) but where incentives are provided by leveraging an agent's reputation concerns as in Holmström (1999). Thus, we revisit a well-known trade-off between incentives - which requires punishing low performance - and exclusion but, in our set-up, we can study vis-à-vis which audience is the exclusion most efficient.

Information design. This paper also borrows from the extensive information design literature. The closest papers to ours are Mukherjee (2008) and Rodina (2016) who find sufficient conditions for full disclosure of the agent's output to be optimal or to maximize the agent's provision of effort in a two-period game and a homogeneous audience in the second period. Relatedly, but with adverse selection in both periods Calzolari and Pavan (2006) finds sufficient conditions for privacy to be optimal. Finally, Ball (2019) study a situation symmetric to ours. In their model, an agent has a multi-dimensional type and a principal who observes signals about each dimension designs a disclosure policy for a single audience. Our work complements theirs by focusing on heterogeneous audiences.

Multiple audiences. Several papers consider set-ups with multiple and (possibly) heterogeneous audiences.

First, our paper also complements a literature on information sharing with heterogeneous audiences who have imperfectly aligned - or even opposed - preferences. Then, an agent building her reputation faces a dilemma because she must choose which audience to impress. For instance, Bouvard and Levy (2018) apply this idea in the context of certification markets where a monopolist's good reputation in accuracy can attract high quality sellers but repel low quality sellers. In companion papers, Bar-Isaac and Deb (2014b) and Bar-Isaac and Deb (2014a) study the role of heterogeneity in audiences' preferences and in audiences' access to information. In the former, they build a general model capturing the various returns to reputation that can arise in various situations where an agent faces audiences with heterogeneous preferences - such as horizontal and vertical competition - and her performance is publicly observable. They show that incentives for effort depends on the curvature of the returns to reputation. In the latter, they focus on an extreme case with two audiences with opposed preferences and show that, while publicly revealing an agent's performance creates countervailing incentives, it allows for compromise while private disclosure leads to suspicion. Therefore, in their set-up, public disclosure dominates differentiated disclosure from a welfare perspective. Compared with this literature, in our model audiences have aligned preferences but there is heterogeneity in information's usefulness. Therefore, there are no countervailing incentives but rather the question is find how to disclose information efficiently.

Second, several papers consider models with multiple audiences with applications to privacy, like ours. Tirole (2021) study situations in which an agent has transient and stable relationships and asks what information to disclose to each audience in order to incentivize altruistic behaviors. However, both audiences value information equally and the agent's has heterogeneous reputation concerns vis-à-vis them. Sharing information more widely boosts incentives but increases ostracism in the private and the public spheres. Relatedly, Frankel and Kartik (2019) show that adding (homogeneous) audiences increases an agent's signaling motives which can worsen the information disclosed in equilibrium because she uses her gaming ability to muddle information. While we also consider multiple audiences, our paper focuses on the heterogeneous value of information for the audiences.

Finally, Horner and Lambert (2016) model a dynamic game with career concerns where a Principal designs a disclosure policy to control what the market learns about the agent's talent. Their models bear similarities with ours in that they allow for differentiated disclosure across time - rather than across heterogeneous employers - and information obtained recently is more informative about the agent's current type - rather than to employers hiring for similar jobs . However because they study learning, the trade-off is dynamic and is between better screening - which requires disclosing more information - and higher incentives - which requires keeping some uncertainty to boosts signal motives.

1.3 The model

1.3.1 Set-up

We consider a two-period model, $t = 1, 2$, with an agent, a firm, and a continuum of potential employers $\rho \in [0, 1]$. The agent may be hired by the firm in period 1, and by one of the employers in period 2. There is ex ante uncertainty regarding which job opportunity may arise and the identity ρ of the employer is only revealed at the beginning of the second period.

There are no transfers and therefore no contracts with financial incentives. However, the firm can leverage the agent's career concerns to provide reputational incentives by strategically disclosing information to the employer. Specifically, at the beginning of the game, the firm publicly commits to a disclosure policy[7]. If the agent accepts the firm's job offer then, after observing her output, the firm sends information to the employer according to the disclosure policy. There are no restrictions on the disclosure policies and therefore the information disclosed can be contingent on the agent's performance as well as on the job opportunity. In the second period, the employer receives the information, updates its beliefs about the agent's talent, and makes a binary hiring decision[8].

We assume that the firm and the employers cannot contract - in particular, the firm cannot sell information to the employer - and therefore the firm does not internalize date-2 profits. Therefore, we refer to this situation as one of private contracting with no information resale.

Agent's talent. The agent has job-specific talents, correlated between each other. Formally, the agent's talent for the date-1 job is $\theta \in \mathbb{R}$, with $E[\theta] = 0$, and is equally unknown to all players. At the beginning of the second period, Nature randomly draws $\rho \in [0, 1]$ according to a distribution with c.d.f G and p.d.f. g, and discloses it publicly. ρ refers interchangeably to the employer or to the job offered by the employer, and it is the correlation between the agent's talents for the date-1 job and for job ρ. Specifically, the agent's talent for job ρ is θ with probability ρ and is randomly drawn from a distribution with mean 0 with probability $1 - \rho$. Thus, given θ, the agent's expected talent for job ρ is $\rho\theta$. The interpretation is that similar jobs require similar skills and therefore correspond to ρ close to 1.[9]

Production Technology. In the first period, the agent's output $y = \theta + e + \varepsilon$ is stochastic and additive[10] in her talent, θ, her effort provision, $e \geq 0$, and noise, ε, satisfying $E[\varepsilon] = 0$ which captures randomness in the outcome as well as measurement errors.

Let F and f denote the c.d.f. and p.d.f. of $\theta + \varepsilon$. We make the following assumptions, summarized below under Assumption 1. The first two assumptions are standard ordering

[7]This commitment assumption can be micro-founded in a repeated cheap talk game where a long-lived principal makes recommendations and needs to have a reputation for reliable communication. See Best and Quigley (2017), Margaria-Smolin (2018) and Mathevet-Pearce-Stacchetti (2019).

[8]The assumption that the employer cannot offer output-contingent contracts is almost with loss of generality. Indeed, if there were agency costs then the employer would want to screen even if it could write contingent contracts.

[9]More generally, to use the jargon of the European Commission in its proposal for an Artificial Intelligence Act, the context in which the information is disclosed is said to be related to the context where it was collected when ρ is close to one and unrelated when ρ is close to 0.

[10]The assumption of additive production technology is canonical and guarantees that, in equilibrium, effort does not affect learning about the agent's talent. It does not affect the results qualitatively - the characterization of efficient disclosure policies does not rely on additivity - but guarantees that learning does not interfere with the trade-offs between incentivizing effort, protecting the agent's reputation and screening talent.

conditions which respectively imply that talent and effort increase the distribution of output in the sense of first order stochastic dominance. While these are natural assumptions, they are not satisfied by all distribution functions, as is well known in the moral hazard literature. The third is a technical assumption which simplifies our computation by allowing us to identify the sign and the monotonicity of several variables.

Assumption 1. *(i)* $E[\theta|y,e]$ *non-decreasing in* y *for all* $e \geq 0$
(ii) $\frac{-f'(y-e)}{f(y-e)}$ *non-decreasing in* y *for all* $e \geq 0$ *(Marginal Likelihood Ratio Property, MLRP)*
(iii) f *is symmetric and single-peaked at* 0

Learning. First, the agent's output is informative about her date-1 talent. While the agent privately chooses effort, her maximization problem is common knowledge and therefore, in equilibrium, all players correctly conjecture her effort provision. Fix the agent's date-1 output y and her date-1 effort provision e^*. We have $E[\theta|y,e^*] = \int \theta f_{\theta|y,e^*}(\theta)d\theta$, where $f_{\theta|y,e^*}(\theta) = \frac{f_\varepsilon(y-e^*-\theta)f_\theta(\theta)}{f_{\varepsilon+\theta}(y-e^*)}$ is a function of $y-e^*$ only. That is, all the useful information about the agent's date-1 talent is contained in $y-e^*$ and, in particular, effort itself does not affect learning.[11]. This observation is important in the model because it guarantees that the firm does not design the disclosure policy so as to learn about the agent's talent but only in order to incentivize effort.

Second, because the agent's talents are correlated, information about her date-1 output is useful to infer her talent in other jobs. In particular, output information is more useful for jobs similar to the date-1 job. Indeed, it is most informative about the agent's talent in job $\rho = 1$. On the contrary, output information is useless for job $\rho = 0$ because it is uninformative about the agent's talent for their job.

Disclosure policy. At the beginning of period 1, the firm publicly commits to a disclosure policy, ψ^*, which maps date-1 output to a probability distribution over some message space. Building on Myerson (1982)'s obedience principle we can, without loss of generality, restrict our attention to incentive compatible direct mechanisms. That is, in the first period, the firm recommends an effort target $e^* \geq 0$ to the agent and, in the second period, it recommends the employer to hire the agent or not by sending positive ($\psi = +$) or negative ($\psi = -$) recommendations. Let $\psi^*_\rho(y)$ denote the probability that the firm sends a positive recommendation $\psi = +$ to employer ρ when the agent's date-1 output is y. In order to guarantee that the firm perfectly controls the disclosure of information, we assume that the agent cannot disclose her

[11]The fact that effort does not affect learning can be achieved with more general production technology: all is needed is separability in effort; ie. $y = h(e)f(\theta,\varepsilon)$. However, we have not yet checked which conditions are then needed for the first-order approach to be valid in this more general set-up.

date-1 output to the employer - neither with hard evidence nor through cheap talk.[12]

1.3.2 Payoffs and strategies

In this subsection we first provide the timing of the game and then set each player's payoffs and strategies.

Timing of the game. In period 1, the firm publicly commits to a disclosure policy, ψ^*, with effort target $e^* \geq 0$. The agent learns ψ^* and either accepts the firm's offer or gets her outside option, which provides her payoff $U_0 \geq 0$. Upon working for the firm, she privately chooses an effort provision $e \geq 0$ at cost $c(e)$ where $c', c'' > 0$ and $c'(0) = 0$. Output y realizes.

In period 2, Nature randomly draws $\rho \in [0, 1]$ and discloses it publicly. After observing ρ and the agent's date-1 output y, the firm sends recommendation $\psi \in \{-, +\}$ according to the disclosure policy ψ^*. employer ρ observes the recommendation ψ, updates its beliefs about the agent's talent for job ρ according to Bayes' rule, and decides to hire the agent or not. The agent receives payoff $u_2 > 0$ if hired and 0 otherwise. The second period output realizes. The game ends.

Employers. In the second period, the employer ρ who was drawn makes a binary decision to hire the agent or not. Because the second period is the last one and there are no transfers, the employer cannot write an explicit incentive contract and therefore the agent exerts no effort. The employer is risk neutral and its profit is affine in the agent's talent. Specifically, hiring the agent for job ρ after receiving recommendation $\psi \in \{-, +\}$ when the effort target is e^* yields expected profit

$$\Pi_\rho^\psi = \underline{\pi} + \rho E[\theta|\psi]$$

where $\underline{\pi} < 0$ is the expected profit from hiring the agent in the absence of information about her date-1 output (because $E[\theta] = 0$). $\underline{\pi}$ is independent of ρ and therefore potential employers only differ by how useful information is to them. Furthermore, $\underline{\pi} < 0$ implies that, by default, the agent is not hired in the second period, and we say that there is severe adverse selection.

Given the disclosure policy ψ^* with effort target e^*, we say that employer ρ is obedient if it always follows the firm's hiring recommendations. For this, it must be profitable to hire the agent if and only if $\psi = +$, which defines two obedience constraints per employer, (IC_ρ^-) and

[12] Otherwise there will be unraveling. In equilibrium, all agents who could be hired if the employer observed their output would disclose it, and the others would not be hired. The situations where this assumption would typically fail are those where the agent has a reputation for honesty or can credibly commit to tell the truth. This could be the case, for instance, in close or repeated relationships such as between friends or siblings.

(IC_ρ^+), which write as follows:

$$\Pi_\rho^- = \underline{\pi} + \rho \frac{\int E[\theta|y,e^*](1-\psi_\rho^*(y))f(y-e^*)dy}{\int(1-\psi_\rho^*(y))f(y-e^*)dy} < 0 \qquad (IC_\rho^-)$$

$$\Pi_\rho^+ = \underline{\pi} + \rho \frac{\int E[\theta|y,e^*]\psi_\rho^*(y)f(y-e^*)dy}{\int \psi_\rho^*(y)f(y-e^*)dy} \geq 0 \qquad (IC_\rho^+)$$

Under severe adverse selection ($\underline{\pi} < 0$) the difficulty for the firm is to persuade the employer to hire the agent. As a result, in equilibrium, (IC_ρ^-) is automatically satisfied and only (IC_ρ^+) needs to be considered.

Overall, given the disclosure policy ψ^* with associated effort target e^*, employer ρ hires the agent if and only if a positive recommendation is sent and thus its ex ante expected profit is

$$\Pi_\rho = \Pi_\rho^+ \int \psi_\rho^*(y)f(y-e^*)dy = \int(\underline{\pi} + \rho E[\theta|y,e^*])\psi_\rho^*(y)f(y-e^*)dy$$

The agent. Upon working for the firm, the agent privately chooses effort $e \geq 0$ to maximize her intertemporal payoff. Her program writes

$$\max_{e\geq 0} \; E[u_2\psi(y)|e] - c(e)$$

We say that the effort target is incentive compatible if, given that employers are obedient, it is in the agent's best interest to exert effort $e = e^*$. Formally, the agent's incentive compatibility constraint, (IC_A), writes

$$e^* \in \arg max_{e\geq 0}\Big\{E[u_2\psi(y)|e] - c(e)\Big\} \qquad (IC_A)$$

Finally, the agent chooses to work for the firm if and only if her expected intertemporal payoff exceeds her reservation utility. This defines the agent's participation constraint

$$E[u_2\psi(y)|e^*] - c(e^*) \geq U_0 \qquad (IR_A)$$

The firm. The firm sets a disclosure policy ψ^* with effort target e^* so as to maximize her profit, equal to date-1 profit because it cannot resell information to the employers. Date-1 profit is equal to the agent's date-1 output, y. Given that the production technology is additive and that the agent's date-1 talent is unknown to all players, the firm's expected profit given the agent's effort provision, $e^* \geq 0$, is $E[y|e^*] = e^*$. Overall, the firm's program writes

$$\max_{\psi^*, e^*} \; e^*$$

$$\text{s.t.} \begin{cases} E_{y,\rho}\left[u_2 \psi^*_\rho(y)|e^*\right] - c(e^*) \geq U_0 & (IR_A) \\ e^* \in \arg max_{e \geq 0} \left\{ E_{y,\rho}[u_2 \psi^*_\rho(y)|e] - c(e) \right\} & (IC_A) \\ \Pi^+_\rho \geq 0 & (IC^+_\rho), \forall \rho \in [0,1] \end{cases}$$

1.4 Single audience

We start by solving the model in the case with a single audience; ie. a single potential employer. Assume that the distribution G is degenerate such that employer ρ is drawn with certainty.

We use a first-order approach, and will show ex post that it is valid under Assumptions 1, 2 and 3, defined at the end of the subsection. Under the first-order approach, we can replace the agent's incentive compatibility constraint by the first-order condition of her maximization problem. Also, it is convenient to simplify (IC^+_ρ) by multiplying each side of the inequality by $\int \psi^*_\rho(y) f(y - e^*) dy$ and regrouping all the terms. Then, the firm's program writes

$$\max_{\psi, e^*} \; e^*$$

$$\text{s.t.} \begin{cases} E\left[u_2 \psi_\rho(y)|e^*\right] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = E\left[u_2 \psi_\rho(y) \frac{(-f'(y-e^*))}{f(y-e^*)}|e^*\right] & (IC_A) \\ E\left[\left(\underline{\pi} + \rho E[\theta|y,e^*]\right)\psi_\rho(y)|e^*\right] \geq 0 & (IC^+_\rho) \end{cases}$$

Let $\lambda \geq 0$ and μ denote the Lagrange multipliers associated with, respectively, the agent's participation constraint (IR_A) and incentive compatibility constraint, (IC_A). Let $\alpha^+_\rho \geq 0$ denote the multiplier associated with the employer's obedience constraint, (IC^+_ρ). The Lagrangian is linear in $\psi_\rho(y)$ and therefore the solution is bang-bang; $\psi_\rho(z) \in \{0, 1\}$ almost everywhere. We derive the first-order conditions characterizing the efficient disclosure policy

$$\psi_\rho(y) = 1 \iff \left(\lambda + \mu \frac{(-f'(y - e^*))}{f(y - e^*)}\right) u_2 + \left(\underline{\pi} + \rho E[\theta|y, e^*]\right)\alpha^+_\rho \geq 0 \qquad (1.1)$$

A standard result in the moral hazard literature is that $\mu > 0$. Its interpretation is that, in equilibrium, the firm would like the agent to exert strictly more effort. Because in this paper we consider situations where the firm wants to encourage rather than discourage effort, we assume that the conditions listed in Assumptions 2 are satisfied, which guarantees that $\mu > 0$. The notation used in Assumptions 2 will be introduced in the discussion below.

Cutoff structure. Given Assumptions 1(i)-(ii) and because all Lagrange multipliers are positive, the left-hand side in Equation (1.1) is non-decreasing in y. As a result, the firm recommends the agent to each employer ρ according to a cutoff rule on her output. That is, for $\rho \in [0,1]$ there exist a cutoff y^*_ρ such that the firm sends a positive recommendation if and only if $y \geq y^*_\rho$.

The intuition for this characterization is two-fold.[13] First, Assumption 1-(i) ensures that the distribution of output increases with talent and therefore a cutoff rule relaxes the employer's obedience constraint (IC^+_ρ), all else equal. Similarly, Assumption 1-(ii) ensures that the distribution of output increases with effort and therefore, given an effort target, a cutoff rule relaxes the agent's participation constraint (IR_A).

Mechanism. Before characterizing the disclosure policy solving the firm's program, let us first discuss the mechanism behind cutoff structures. Consider a disclosure policy recommending the agent according to a cutoff rule and let y^* denote the cutoff.

First, from the agent's incentive compatibility constraint, we find that the effort provision e^* associated with cutoff y^* satisfies $c'(e^*) = u_2 f(y^* - e^*)$. Taking the derivative with respect to y^*, and given that f is single-peaked at 0, we find that effort - and hence, the firm's expected profit - is hump-shaped in the cutoff. In particular, e^* increases with y^* as long as $y^* - e^* \leq 0$ and achieves its maximum, $\overline{e}$, when the cutoff is $\overline{y} = \overline{e}$.

Second, consider $e^*(y^*)$ satisfying the agent's incentive compatibility constraint and suppose that $y^* - e^* \leq 0$. The agent's intertemporal payoff is $(1 - F(y^* - e^*))u_2 - c(e^*)$ and strictly decreases with the cutoff y^*[14]. The intuition is that a higher cutoff leads to higher effort, which is costly, and to more exclusion in the second period. Thus, the disclosure policy maximizing the agent's intertemporal payoff induces effort $\underline{e}$ and is implemented with the lowest cutoff satisfying the employer's obedience constraint, $\underline{y} = y^+_\rho(\underline{e})$. Note that, because adverse selection is severe ($\underline{\pi} < 0$), $\underline{y}$ is finite and therefore complete privacy is not optimal for the agent. Furthermore, in equilibrium, the firm sets a cut-off y^* associated with effort e^* such that $y^* - e^* \leq 0$. Otherwise, if $y^* - e^* > 0$ then the firm can induce strictly higher effort by setting a lower cutoff.

Third, employer ρ's ex ante expected profit is $\Pi_\rho = \int_{y^*} (\underline{\pi} + \rho E[\theta|y, e^*]) f(y - e^*) dy$. Fixing

[13] Although it incorporates an endogenous type (due to moral hazard), the logic is the same as in Kamenica and Gentzkow (2011) example B on supplying product information, Lerner and Tirole (2006), and a number of other papers.

[14] We have that e^* strictly increases with y^* as long as $y^* - e^* \leq 0$. Furthermore, from (IC_A) it is straightforward to find that $y^* - e^*$ strictly increase with y^* for all $y^* - e^* \in \mathbb{R}$.

the agent's effort provision, e^*, the employer's expected profit is hump-shaped in the cutoff y^*.[15] Let $y_\rho^{**}(e^*)$ denote the cutoff maximizing employer ρ's expected profit. By definition, $\underline{\pi} + \rho E[\theta|y = y_\rho^{**}, e^*] = 0$ and, because $\underline{\pi} < 0$ and $E[\theta|y = y_\rho^{**}, e^*]$ has same sign as $y_\rho^{**} - e^*$, then $y_\rho^{**} - e^* > 0$. The disclosure policy maximizing the employer's profit perfectly discloses the agent's output. Under full transparency the effort is $\bar{\bar{e}}$ and employer ρ hires the agent if and only if her output exceeds $\bar{\bar{y}} = y_\rho^{**}(\bar{\bar{e}})$. However recall that, in equilibrium, $y^* - e^* \leq 0$ which means that it is never optimal for the firm to be be fully transparent.

To sum up, when designing a disclosure policy, the firm faces a trade-off between incentives - which involves punishing low output - and participation - which requires providing sufficiently high payoff. This trade-off materializes in the setting of a cutoff on the agent's output. A higher cutoff incentivizes higher effort, which increases the firm's profits, but also decreases the agent's payoff. Incidentally, increasing the cut-off also increases the employer's profit. However, the objective of the firm is not to increase the employer's profit but to exploit its willingness to screen talent in order to manipulate its hiring decision, which indirectly incentivizes effort. We illustrate this mechanism on Figure 1 where we represent the agent's equilibrium payoff and the firm's equilibrium profit achieved with cutoff rules by varying the cutoff from $\underline{y}$, the lowest one satisfying the employer's obedience constraint, to $+\infty$.

We summarize the characterization of the disclosure policy solving the firm's program in the case of private contracting with no information resale in Proposition 1.

Proposition 1. *(Private contracting with no information resale - single audience)*
Under private contracting with no information resale the firm's objective is to maximize the agent's effort. Under Assumptions 1, 2 and 3, the disclosure policy ψ^* solving the firm's program has a cutoff structure. A higher cutoff incentivizes higher effort but yields lower payoff to the agent. The optimal cutoff y_ρ^* is the highest cutoff in $[\underline{y}, \overline{y}]$ satisfying the agent's participation constraint.

Proof. See Appendix 1.A.1. □

Sufficient conditions. We now present sufficient conditions for the existence of solutions to the firm's program and for the first-order approach to be valid.

Based on the discussion above, for the existence of solutions to the firm's program, with $\mu > 0$, it is necessary (and sufficient) that there exist (y^*, e^*) with $y^* - e^* \leq 0$ and satisfying (IR_A), (IC_A) and (IC_ρ^+). These conditions satisfied under the following sufficient conditions listed in Assumption 2.

[15] Mathematically, this comes from the fact that $E[\theta|y, e^*]$ increases with y (Assumption 1(i)) and has same sign as $y - e^*$ (Assumption 1(i)).

Assumption 2. *(Existence of solutions)*

(i) $y_\rho^+(\underline{e}) - \underline{e} < 0$

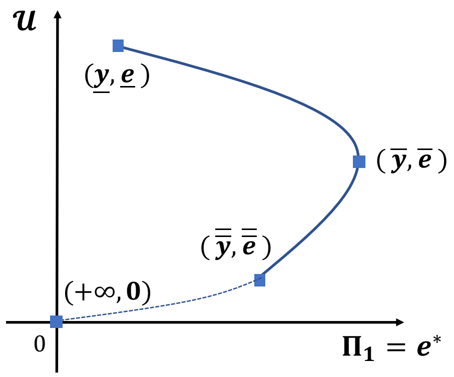
(ii) $\big(1 - F(\underline{y} - \underline{e})\big)u_2 - c(\underline{e}) \geq U_0$

While we can show that under Assumptions 1 and 2, e^* is a local maximum in the agent's program, it may have a global maximum at some effort strictly below e^*. The following assumption guarantees that, given any cutoff, the agent's problem is single peaked.

Assumption 3. *(First-order Approach)*

Let $\overline{y}$ denote the cut-off implementing the highest effort $\overline{e}$; $\overline{y} = \overline{e}$. Then, $f(\overline{y} - e)u_2 - c'(e) \geq 0$, for all $e \in [0, \overline{e}]$

Figure 1 – Equilibrium payoffs attainable for the agent and the firm with cutoff rules satisfying the agent's incentive compatibilty constraint (IC_A)



Discussion. Let us now interpret in words the mechanism of reputational incentives behind the disclosure policies characterized in Proposition 1. Because disclosure policies are cutoff rules we cannot use classic notions of informativeness - such as Blackwell ranking - to measure how much information is disclosed to the employer. Instead, we use the employer's expected profit as a proxy for informativeness; see Definition 1, below. This is a natural proxy because the employer does not value information itself, but only the profit it derives from using it. Furthermore, we have shown that the employer's expected profit strictly increases with the cutoff and therefore this measure of informativeness is well defined.

Definition 1. *(Informativeness)*

A disclosure policy ψ^* is more informative than $\psi^\dagger$ if and only if the employer's expected profit is strictly greater under ψ^*.

Under this definition, the mechanism can be interpreted as follows. In order to incentivize effort, the firm discloses information about the agent's output to the future potential employer.

Disclosing more information allows the employers to better screen talent and, in turn, the agent increases effort to avoid higher risk of unemployment.

1.5 Heterogeneous audiences

We now turn to the case with heterogeneous audiences who differ in how useful information is to them and thus, we no longer assume that G is degenerate. In the baseline model with a single audience there was no distinction between private and public disclosure of information. The goal now is to understand how heterogeneity affects the disclosure of information and, in particular, when to disclose differentiated information to the audiences.

1.5.1 Benchmark with contractible effort

Consider, as a benchmark, the case where date-1 effort is contractible through some unspecified mechanism and where the disclosure policy is designed to maximize the agent's continuation payoff.[16] Let $e^B > 0$ denote this effort. Then, only the employers' obedience constraints are relevant and the program writes

$$\max_{\psi} E_{y,\rho}\left[u_2 \psi_\rho(y) | e^B\right]$$
$$\text{s.t. } \Pi_\rho^+ \geq 0 \qquad\qquad (IC_\rho^+), \forall \rho \in [0,1]$$

As in the baseline model with a single audience, we will solve the program with a first-order approach. In this program, no assumptions are required except that functions are differentiable.

Let ψ^B denote the benchmark disclosure policy, solution to the above program. As in the case with a single audience, ψ^B has a cutoff structure and the agent is recommended to each employer ρ according to a cut-off rule on output. Furthermore, at the optimum, all employers' incentive constraints (IC_ρ^+) are binding. Otherwise, the agent could be recommended strictly more often to some employers and still be hired, which would strictly increase her payoff. Let $y_\rho^+(e^B)$ denote the cutoff satisfying the employer's obedience constraint for job ρ when date-1 effort is e^B. On Figure 2 we represent the form of the disclosure policy maximizing the agent's continuation payoff.

[16]This benchmark is also relevant from a policy perspective. For instance, in its proposal for an Artificial Intelligence Act, the European Commission's goal is to ensure that AI is "human-centric".

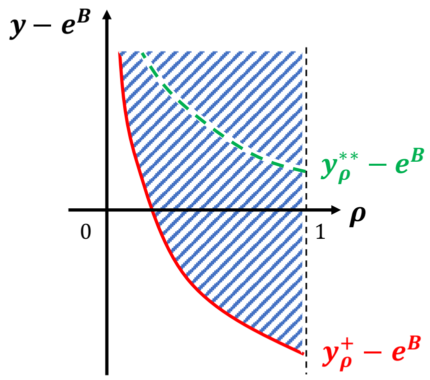


Figure 2 – Benchmark disclosure policy, ψ^B, maximizing the agent's continuation payoff.
Note. Given the agent's effort provision e^B, the employer ρ's obedience constraint binds at the cut-off $y_\rho^+(e^B)$, which is represented in a red line on the Figure. The employer ρ's expected profit is maximal when the cut-off is $y_\rho^{**}(e^B)$, which is represented in a green dashed line on the Figure.

Discussion. We now discuss the form of the benchmark disclosure policy.

First, complete opacity about the agent's past performance cannot be optimal because, otherwise, she is never hired. Thus, except for employer $\rho = 0$ for which information is useless, some information is disclosed to all employers and therefore the cut-off is finite for all jobs.

Second, the cutoff $y_\rho^+(e^B)$ strictly decreases with ρ which means that the agent is hired more often for jobs similar to her date-1 job. The reason is that it is easier to persuade the employer to hire the agent when holding useful information about her talent. Indeed, upon receiving a positive recommendation $\psi = +$, hiring the agent for job ρ yields expected profit $\underline{\pi} + \rho E[\theta|\psi = +, e^B]$, with $\underline{\pi} < 0$. If it is profitable to hire the agent for job ρ, then $E[\theta|\psi = +, e^B] > 0$ and therefore it is also profitable to hire the agent for job $\rho' \geq \rho$. That is, although all employers have the same prior about the the agent, high-ρ employers require less optimistic signals about the agent's date-1 talent to decide to hire her.

Third, public disclosure policies cannot be optimal for the agent because the benchmark disclosure policy cannot be implemented with public messages. The reasoning builds on the one above. Under ψ^B, when a positive recommendation is sent to employer ρ', the signal is insufficiently optimistic about the agent's date-1 talent for all employers $\rho < \rho'$. Therefore, if ψ^B were implementable with public messages, then the agent would never be hired by any employer $\rho < \rho'$ following any output $y \geq y_\rho^+(e^B)$.

Reputation protection. In what follows we study the disclosure policies designed for various objectives and it will be useful to interpret their characterizations as deviations from the benchmark disclosure policy maximizing the agent's hiring rate. Because disclosing information about the agent's output affects her reputation - and, therefore, her chances to be hired - we introduce the following definitions about protecting and jeopardizing reputation.

Definition 2. *(Reputation protection)*
We say that the agent's reputation vis-à-vis employer ρ is protected if and only if its obedience constraint binds. Otherwise, we say that her reputation is jeopardized.

1.5.2 Private contracting with no information resale

We now study the case of private contracting with no information resale, as in the baseline model, but with heterogeneous audiences. The firm's program writes

$$\max_{\psi, e^*} \; e^*$$

$$\text{s.t.} \begin{cases} E_{y,\rho}\left[u_2 \psi_\rho(y)|e^*\right] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = E_{y,\rho}\left[u_2 \psi_\rho(y) \frac{(-f'(y-e^*))}{f(y-e^*)}|e^*\right]\right] & (IC_A) \\ E\left[\left(\underline{\pi} + \rho E[\theta|y]\right)\psi_\rho(y)|e^*\right] \geq 0 & (IC_\rho^+), \forall \rho \in [0,1] \end{cases}$$

As in the baseline model, we solve the problem with a first-order approach. It is valid under Assumptions 4 and 5, which are the equivalent of Assumptions 2 3 in the case with heterogeneous audiences.

Assumption 4. *(Existence of solutions)*
(i) $y_\rho^+(\underline{e}) - \underline{e} < 0$
(ii) $E_\rho\left[1 - F(y_\rho^+(\underline{e}) - \underline{e})\right]u_2 - c(\underline{e}) \geq U_0$

Assumption 5. *(First-order Approach)*
Let $\overline{y}$ denote the cut-off implementing the highest effort $\overline{e}$. Then, $E_\rho[f(\max\{\overline{y}, y_\rho^+(\overline{e})\} - e)u_2] - c'(e) \geq 0$, for all $e \in [0, \overline{e}]$

As in the baseline model, under Assumption 4 $\mu > 0$ and thus all Lagrange multipliers are positive. Therefore the disclosure policy is a cut-off rule on the agent's date-1 output and the efficient cutoff for job ρ, satisfies Equation (1.2) with equality.

$$\psi_\rho(y) = 1 \iff \left(\lambda + \mu \frac{(-f'(y-e^*))}{f(y-e^*)}\right)u_2 + \left(\underline{\pi} + \rho E[\theta|y]\right)\alpha_\rho^+ \geq 0 \tag{1.2}$$

There are two cases to consider. If the employer ρ's obedience constraint binds then the cutoff is $y_\rho^+(e^*)$. If the employer ρ's obedience constraint is slack then $\alpha_\rho^+ = 0$ and the cutoff solves

$$\lambda + \mu \frac{-f'(y-e^*)}{f(y-e^*)} = 0 \tag{1.3}$$

Equation 1.3 is independent of ρ and therefore the cut-off is uniform across employers whose obedience constraints are slack. Let y^* denote this uniform cutoff. We have that, for each $\rho \in [0,1]$, the cutoff is $\max\{y^+_\rho(e^*), y^*\}$. Furthermore, since $\lambda \geq 0$ and $\mu > 0$ we get $y^* - e^* < 0$. Let us now characterize the set of implementable uniform cutoff - to ease notation, we use the same one as in the case with a single audience.

First, we have seen in the benchmark case that it is easier to recommend the agent to employers hiring jobs similar to the date-1 job. Therefore, the lowest implementable cutoff is such that employer $\rho = 1$'s obedience constraint bind; $\underline{y} = y^+_{rho=1}(\overline{e})$ where $\underline{e}$ denote the agent's effort when all obedience constraints bind and

Furthermore, by taking the derivative of the agent's incentive compatibility constraint (IC_A) with respect to y^* we find that e^* strictly increases with y^* as long as $y^* - e^* < 0$. Thus, as in the baseline model with a single audience, effort is highest when the cutoff is $\overline{y} = \overline{e}$.

We summarize this finding in Proposition 2 and represent an example of an efficient disclosure policy designed under private contracting with no resale on Figure 3.

Proposition 2. *(Private contracting with no information resale - heterogeneous audiences)*
Under private contracting with no information resale the firm's objective is to maximize the agent's effort. Under Assumptions 1, 4 and 5, the disclosure policy ψ^* solving the firm's program has a cutoff structure and the cutoff is uniform across jobs where the employer's obedience constraint is slack.

Under this cutoff structure, the agent's reputation is jeopardized vis-à-vis employers hiring for jobs similar to the date-1 job and protected vis-à-vis the other employers. A higher uniform cutoff jeopardizes the agent's reputation vis-à-vis more employers and discloses more information to them. These two channels incentivize higher effort but yield lower payoff to the agent. The optimal uniform cutoff y^* is the highest cutoff in $[\underline{y}, \overline{y}]$ satisfying the agent's participation constraint.

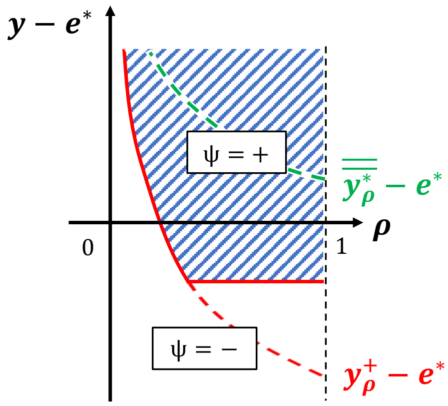


Figure 3 – Example of a disclosure policy under private contracting with no information resale. The cutoff set by the firm is represented in the red full line.
Note. Given the agent's effort provision e^*, the employer's obedience constraint for job ρ binds at the cut-off $y^+_\rho(e^*)$, which is represented in a red dotted line on the Figure. The employer's expected profit for job ρ is maximal when the cut-off is $\bar{\bar{y}}^*_\rho(e^*)$, which is represented in a green dashed line on the Figure.

Discussion. We now compare the disclosure policies in the heterogeneous audiences and single audience cases in order to highlight the specific features stemming from audience heterogeneity.

The mechanism of reputational incentives is similar in both cases. Indeed, in the case with heterogeneous audiences, the firm sets a uniform cutoff rather than a single cut-off. Increasing the cutoff increases the risk of exclusion and therefore incentivizes effort. However, there are important distinctions which are specific to heterogeneous audiences.

First, the cutoffs are differentiated and the firm jeopardizes the agent's reputation vis-à-vis employers hiring for jobs similar to the date-1 job but protects it vis-à-vis the other employers. The reason is that a uniform cutoff would violate the obedience constraint of the employers hiring for the jobs most different to the date-1 job (ie. ρ low). Therefore, when increasing the uniform cutoff, the firm not only discloses more information to the employers (as in the single audience case) but he also jeopardizes the agent's reputation vis-à-vis more employers, from the ones most interested in output information to the least interested ones.

Second, the firm sends differentiated recommendations only to protect the agent's reputation. employers vis-à-vis which the firm jeopardizes the agent's reputation receive the same information - because the cut-off is uniform. Furthermore, the cutoffs for the other employers are all higher than the uniform cut-off, which means that the uniform cut-off does not affect their hiring decisions. Therefore, the disclosure policy under heterogeneous audiences could be implemented semi-publicly. The firm can send public recommendations according to a cutoff rule with the uniform cutoff and, on top of this, send private recommendations to protect the agent's reputation vis-à-vis the employers hiring for the jobs least related to the date-1 job.

1.6 Stake in date 2 / Internalizing date-2 profits

Up to now, we have assumed that the firm has no stake in the second period. As a result, the firm designs a disclosure policy only to incentivize the agent to exert effort and typically does not internalize the effect of disclosing information on date-2 profits. However, there are situations in which these externalities are internalized, for instance by a social planner maximizing welfare or when there are vertical relationships such as information resale or when firms are integrated (eg. multi-sided platforms) or if the firm can re-hire the agent. In this section, we will study such situations.

1.6.1 Re-selling information to a single employer

Assume, first, that the firm puts a weight 1 on date-1 profits and w_F on the profits of some employer $\rho_F \in [0,1]$. This could be because the firm can resale information to employer ρ_F and has bargaining power w_F or because it may have the opportunity to hire the agent in date 2 for job ρ_F and puts a weight w_F on date-2 profit.

To model this situation, consider the following game, slightly modified from the original one studied earlier. Fix $\rho_F \in [0,1]$. In the second period, with probability γ, employer ρ_F is drawn and, with probability $1-\gamma$, an employer $\rho \in [0,1]$ is randomly drawn according to the distribution G.

The firm now designs two disclosure policies: ψ_F for employer ρ_F and ψ for employers $\rho \in [0,1]$. The firm's program writes

$$\max_{\psi_F,\psi,e^*} \; e^* + \gamma w_F E_y\left[\psi_F(y)(\underline{\pi} + \rho_F E[\theta|y])\right]$$

$$\text{s.t. } (IR_A),(IC_A),(IC^+_{\rho_F}),(IC^+_{\rho}),\forall\rho\in[0,1]$$

The first order conditions are

$$\psi_\rho(y) = 1 \iff \left(\lambda + \mu\frac{-f'(y-e^*)}{f(y-e^*)}\right)u_2 + \left(\underline{\pi} + \rho E[\theta|y]\right)\alpha^+_\rho \geq 0 \tag{1.4}$$

$$\psi_F(y) = 1 \iff \left(\lambda + \mu\frac{-f'(y-e^*)}{f(y-e^*)}\right)u_2 + \left(\underline{\pi} + \rho_F E[\theta|y]\right)(w_F + \alpha^+_{\rho_F}) \geq 0 \tag{1.5}$$

Equation 1.4 is identical to Equation 1.2 and therefore the disclosure policy ψ designed for the employers who are unrelated to the firm has the same structure as in the case of private contracting with no information resale (see Proposition 2).

All terms in Equation 1.5 are non-decreasing in y and therefore ψ_F has a cut-off structure as well: the firm sends a positive recommendation to employer ρ_F if and only if the agent's performance exceeds some cut-off.

It is straightforward to show that the cutoff for employer ρ_F is greater under ψ_F than under ψ; that is, the firm provides more information to employer ρ_F because it internalizes its profits. To see this, recall that under ψ the cut-off $y^* < 0$ is uniform across jobs where the employer's obedience constraint is slack and assume that employer $\rho = \rho_F$'s obedience constraint is slack. Then, we have $\lambda + \mu \frac{-f'(y^*-e^*)}{f(y*-e^*)} = 0$ and, since $y^* < 0$, $\underline{\pi} + \rho_F E[\theta|y] < 0$. Therefore the condition of Eqution 1.5 is not satisfied at $y = y^*$ and $\psi_F(y^*) = 0$.

Finally, internalizing the profit of employer ρ_F does not affect the structure of the disclosure policies to the other employers but it affects how much information is disclosed to them - ie. the level of the uniform cutoff. Indeed, the agent's participation constraint (IR_A) binds in equilibrium whether the firm internalizes employer ρ_F's profit or not. However, when the firm internalizing employer ρ_F's profit, it discloses more information to employer ρ_F which affects the agent's payoff and her incentives to exert effort in the first period. Therefore, it must compensate by disclosing less information to the other employers.

Proposition 3. *(Information resale to a single employer)*
Suppose that the firm puts weight $w_F > 0$ on an employer's profit, hiring for job ρ_F. The firm provides more information to this employer than to other employers offering similar jobs. In order to satisfy the agent's participation constraint, the firm balances providing more information to this employer to increase its profit and protecting the agent's reputation vis-à-vis the other employers. In this trade-off, the disclosure policy for the other employers has the same structure as in the case of private contracting with no information resale, as characterized in Proposition 2.

1.6.2 Re-selling information to all employers

In the previous sub-section we have shown that the firm would provide more information to a date-2 employer than to the others if it can extract the associated rent. We now consider a situation where the firm can contract with all date-2 employers and ask how it then chooses to disclose information.

Suppose that the firm puts no weight on the agent's payoff, a weight equal to 1 on date-1 profit and equal to $w_2 \in (0, 1]$ for all date-2 employers' profits. We will refer to this case as one of private contracting between the firm and the agent with information resale where the firm's bargaining power in the downstream market is w_2. But this set-up also captures a situation of a large integrated firm where employees may be hired by various departments or one where a multi-sided platform guarantees the participation of one side (the agent's side) and maximizes a weighted average of the welfare on the other sides (the firm and the employers' sides).

The firm's program writes

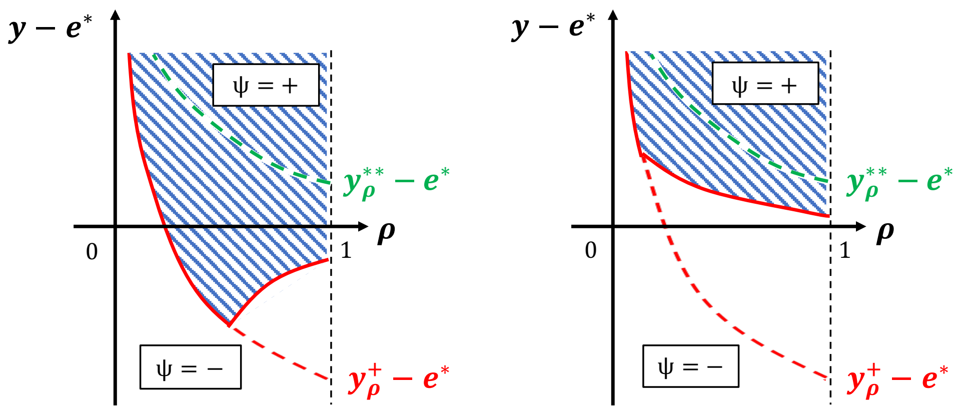
$$\max_{\psi^*, e^*} e^* + w_2 E_{y,\rho}\left[\psi^*_\rho(y)\Pi^+_\rho\right]$$

$$\text{s.t.} \begin{cases} E_{y,\rho}\left[u_2\psi_\rho(y)|e^*\right] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = E_{y,\rho}\left[u_2\psi_\rho(y)\frac{(-f'(y-e^*))}{f(y-e^*)}|e^*\right] & (IC_A) \\ \Pi^+_\rho \geq 0 & (IC^+_\rho), \forall \rho \in [0,1] \end{cases}$$

Here again we use a first-order approach. For each employer ρ, the efficient disclosure policy is a cutoff rule, and the optimal cutoff is $\max\{y^+_\rho(e^*), y^*_\rho\}$, where $y^+_\rho(e^*)$ is the cutoff when employer ρ's obedience constraint binds and y^*_ρ solves the following first-order condition

$$\left(\lambda + \mu\frac{-f'(y-e^*)}{f(y-e^*)}\right)u_2 + \left(\underline{\pi} + \rho E[\theta|y]\right)w_2 = 0 \tag{1.6}$$

Taking the total derivative of the left-hand-side of Equation 1.6 with respect to ρ, we find that $\frac{dy^*_\rho}{de^*} > 0$ if and only if $y^*_\rho - e^* < 0$. Thus, there are two regimes, depending on whether (date-1) incentives motives or (date-2) screening motives dominate, and we represent them on the two panels of Figure 4.



(a) Incentive motives dominate. (b) Screening motives dominate.

Figure 4 – Examples of a disclosure policy under private contracting with information resale downstream in the regime where incentive motives dominate (left) and where screening motives dominate (right).

The regime where screening motives dominate is inefficient from the perspective of the agent and the firm because there exist disclosure policies inducing higher effort and yielding higher payoff for the agent. Therefore, while both regimes are theoretically plausible, in the practical examples we have in mind - such as social scores or blacklists - the central motivation is to encourage effort, not to help employers screen types. For instance, the law requires firms to erase the name of consumers of blacklists after they have paid their due in order to protect their reputation and avoid ostracism, even though the information might be relevant.

This justifies that we rule out the regime where screening motives dominate. This would be the case if w_2 the weight put on date-2 profits is low enough or a disclosure policy under the regime where screening motives dominate would violate the agent's participation constraint.[17] For simplicity, let us simply assume that w_2 is low enough.

We now compare the form of disclosure policies in the cases of private contracting between the firm and the agent with and without information resale. First, the agent's participation constraint (IR_A) binds in each case and therefore the agent is indifferent between both cases.[18] Thus, given that in the case without information resale the objective is too maximize effort, the agent's effort is lower in the case with information resale.

Second, allowing for information resale also changes the shape of the disclosure policy. While in the case without information resale the cutoff was uniform across employers vis-à-vis which the agent's reputation was jeopardized, the cutoff increases with ρ in the case with information resale. The intuition is the following. Recall that the efficient way to incentivize effort is through a cutoff contingent on performance only; ie. a uniform cut-off. Now that the principal accounts for date-2 profits, he must increase some, if not all, cutoffs. The choice of the cutoffs to increase has no effect on the agent's payoff nor on her incentives However, providing more information to high-ρ strictly increases date-2 expected profits.

In Definition 3 we build on our definition of informativeness to compare how much information two different employers get. The result then has the interpretation that, when it internalizes date-2 profits, the firm provides more information to employers hiring for the jobs most similar to the date-1 job. We summarize the results below in Proposition 4.

Definition 3. *(Informativeness comparison)*
Consider a disclosure policy ψ^* and two employers ρ and ρ' vis-à-vis which the agent's reputation is jeopardized. We say that employer ρ is more informed than employer ρ' if and only if both employers strictly prefer to receive the information disclosed to employer ρ.

[17] In the regime where screening motives dominate, increasing any cutoff decreases the agent's effort and decreases her hiring rate. Therefore, it is generally ambiguous whether this increases or decreases her intertemporal payoff.

[18] In the regime where incentive motives dominate, the agent's participation constraint binds. Otherwise, the firm could increase any cutoff, which would increase effort and improve screening.

Proposition 4. *(Information resale to all employers)*
Under private contracting with information resale the firm's objective is to maximize a weighted average between date-1 and date-2 profits. Under Assumptions 1, 2 and 3, the disclosure policy solving the firm's program has a cutoff structure and and the cutoff increases with ρ among employers whose obedience constraint is slack.

Under this cutoff structure, the agent's reputation is jeopardized vis-à-vis employers hiring for the jobs most similar to the date-1 job. Among these employers the firm discloses more information to those who value it the most. This differentiated information disclosure dampens incentives but improves screening.

Discussion. When date-2 profits are internalized the information disclosed is tailored to the audience and therefore the disclosure policy cannot be implemented with public messages. Indeed, if information were publicly disclosed, all audiences would receive the same information. In particular, audiences who value the information the most receive more information. This feature is in line with the observation that several laws compartmentalize individual's information, and provide more information to the parties that need it the most.

For instance, in Europe and in the USA, many employers are not allowed access to personal information about their employees or about job candidates - such as criminal records, credit scores or profiles on social media - with exceptions for sensitive jobs such as police forces or accountants with higher risks of corruption. Similarly, in France, there exist three versions of each criminal record, with more or less crimes listed, and employers' access is regulated. Finally, in Europe, the firms within the same sector of activity can share a blacklists of fraudulent customers, but they are not allowed to share it with the firm from a different sector of activity.

1.6.3 Re-hiring

So far, we have considered situations where, in the second period, the agent only interacts with a single employer - potentially the firm. We now consider situations where, in the second period, both the firm and a randomly drawn employer may hire the agent and study how the firm optimally uses the information she collected in date 1. In particular, we are interested in the difference between the situations where the firm competes with employers who value the information more or less than itself.

The timing of the second period is as follows. First, Nature randomly draws $(\rho_E, \rho_F) \in [0,1]^2$, according to some c.d.f. G_2, the jobs for which the agent can be hired by, respectively, a date-2 employer and the firm. Then, based on the information disclosed by the firm, employer

ρ_E decides whether to hire the agent or not. Finally, if the agent has not been hired, the firm decides whether to hire her or not for some job ρ_F, where $\rho_F \in (0,1)$ is given and publicly known at the beginning of the game.

It is useful to interpret the game as one where, in period 1, the firm discloses information to a date-2 employer and its future self and to study who receives more information. At the beginning of period 1, the firm commits to two disclosure policies, ψ, for the other employers, and ϕ, for itself. Because the firm discloses information privately and the hiring decisions are binary we can, without loss of generality, restrict attention to binary recommendations. Let $\psi_{(\rho_E,\rho_F)}(y)$ (resp. $\phi_{(\rho_E,\rho_F)}(y)$) denote the probabilities to send a positive recommendation to the date-2 employer (resp. the firm) when the agent's date-1 output is y.

In the second period, the firm can sell the information - disclosed according to ψ - to the other employers and we assume that it has bargaining power equal to $w_F \in [0,1]$, uniform across all employers. Let $\pi_E(y) = \underline{\pi} + \rho_E E[\theta|y]$ (resp. $\pi_F(y) = \underline{\pi} + \rho_F E[\theta|y]$) denote the expected profit from hiring the agent for job ρ_E (resp. ρ_F) following output y. When the date-1 output is y and the tasks drawn are (ρ_E, ρ_F), the firm's date-2 expected profit is $w_F\psi_{(\rho_E,\rho_F)}(y)\pi_E(y) + (1-\psi_{(\rho_E,\rho_F)}(y))\phi_{(\rho_E,\rho_F)}(y)\pi_F(y)$.

Finally, to be consistent with the assumption that the firm has commitment power, we impose that ϕ is designed so that the firm's date-2 profits are positive on expectation, for each ρ_F.

The firm's program writes.

$$\max_{\psi,\phi,e^*} e^* + \iint_{(\rho_E,\rho_F)} \int_y \left[w_F\psi_{(\rho_E,\rho_F)}(y)\pi_E(y) + (1-\psi_{(\rho_E,\rho_F)}(y))\phi_{(\rho_E,\rho_F)}(y)\pi_F(y)\right] f(y-e^*) g_2(\rho_E,\rho_F) dy d\rho_E d\rho_F$$

$$\text{s.t.} \begin{cases} u_2 E_{y,\rho_E,\rho_F}\left[\left[\psi_{(\rho_E,\rho_F)}(y) + (1-\psi_{(\rho_E,\rho_F)}(y))\phi_{(\rho_E,\rho_F)}(y)\right]|e^*\right] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = u_2 E_{y,\rho}\left[\left[\psi_{(\rho_E,\rho_F)}(y) + (1-\psi_{(\rho_E,\rho_F)}(y))\phi_{(\rho_E,\rho_F)}(y)\right]\frac{(-f'(y-e^*))}{f(y-e^*)}|e^*\right] & (IC_A) \\ E\left[(\underline{\pi} + \rho_E E[\theta|y])\psi_{(\rho_E,\rho_F)}(y)|e^*\right] \geq 0 & (IC^+_{\psi,\rho_E,\rho_F}) \\ E\left[(\underline{\pi} + \rho_F E[\theta|y])\phi_{(\rho_E,\rho_F)}(y)|e^*\right] \geq 0 & (IC^+_{\phi,\rho_E,\rho_F}) \end{cases}$$

We are unable to provide a complete characterization of the solution of this program in the general case without making additional assumptions on the distribution of talents and the production technology. Instead, we can derive some relevant properties of the efficient disclosure policies, both in the general case and in some special cases.

Proposition 5. *(i) In the second period, for any realization of $(\rho_E, \rho_F) \in [0,1]^2$, the agent is hired according to a cut-off rule on her date-1 performance.*

(ii) When the information is more useful to the firm than to the other employer - ie. when $\rho_F \geq \rho_E$ -

then the firm does not disclose information to the other employer.

(iii) Assume that the firm has bargaining power $w_F = 1$. Then, information is used efficiently in the sense that after each draw (ρ_E, ρ_F) the player who values information the most - ie. $\max\{\rho_E, \rho_F\}$ *- is the only one to be informed.*

(iv) Assume that the firm has no bargaining power $w_F = 0$. Fix any $\rho_E \in [0,1]$. In date 2, the firm recommends the agent to employer ρ_E with stricty positive probability.

Proof. (i) We use a first-order approach. Fix (ρ_E, ρ_F). We have that $\phi_{(\rho_E,\rho_F)}(y) = 1$ if and only if $\left(\lambda + \mu \frac{-f'(y-e^*)}{f(y-e^*)}\right)u_2 + \pi_F(y)\alpha^+_{\psi,\rho_E,\rho_F} \geq 0$. The LHS is non-decreasing in y and therefore $\phi_{(\rho_E,\rho_F)}$ is a cut-off rule. Let $y^*_{(\phi,\rho_E,\rho_F)}$ denote the corresponding cut-off.

For $y \leq y^*_{(\rho_E,\rho_F)}$, we have that $\psi_{(\rho_E,\rho_F)}(y) = 1$ if and only if $\left(\lambda + \mu \frac{-f'(y-e^*)}{f(y-e^*)}\right)u_2 + \pi_E(y)(w_F + \alpha^+_{\psi,\rho_E,\rho_F}) \geq 0$. The LHS is non-decreasing in y. Let $y^*_{(\psi,\rho_E,\rho_F)}$ denote the lowest output such that $\psi_{(\rho_E,\rho_F)}(y) = 1$. Then, in the second period, the agent is hired if and only if $y \geq \min\{y^*_{(\psi,\rho_E,\rho_F)}, y^*_{(\phi,\rho_E,\rho_F)}\}$.

Note that for $y > y^*_{(\rho_E,\rho_F)}$, we have that $\psi_{(\rho_E,\rho_F)}(y) = 1$ if and only if $\pi_E(y)(w_F + \alpha^+_{\psi,\rho_E,\rho_F}) - \pi_F(y) \geq 0$. Wihtout further assumptions, the monotonicity of the LHS is unknown, and therefore we cannot conclude when the agent is hired by the firm or by employer ρ_E.

(ii) Suppose that $\rho_F \geq \rho_E$ and assume, by contradiction, that the firm sometimes sends positive hiring recommendation $\psi = +$ to employer ρ_E. The disclosure policy satisfies employer ρ_E's incentive constraint and thus $\underline{\pi} + \rho_E E[\theta|\psi = +] \geq 0$. Since $\underline{\pi} < 0$, then $E[\theta|\psi = +] > 0$. Therefore the firm's profit would have been $\underline{\pi} + \rho_F E[\theta|\psi = +] > \underline{\pi} + \rho_E E[\theta|\psi = +] > (\underline{\pi} + \rho_E E[\theta|\psi = +])w_F$ if it had re-hired the agent itself, which contradicts the optimality of the policy.

(iii) The same reasoning applies in the case where $\rho_F < \rho_E$ when the firm's bargaining power is 1. Employer ρ_E would make higher profit than the firm if the information it receives were sent according to ϕ as well. Thus. instead of hiring the agent itself, the firm could let employer ρ_E hire her and extract its rent by posting a price equal to its expected profit.

(iv) Suppose that the firm has no bargaining power; $w_F = 0$ and fix some $\rho_E \in [0,1]$.

Let $y^*_{(\phi,\rho_E,\rho_F)}$ denote the cut-off above which $\phi_{(\rho_E,\rho_F)}(y) = 1$. We have shown in the proof of (i) that the firm sends a positive recommendation to employer ρ_E following some output $y < y^*_{(\phi,\rho_E,\rho_F)}$ if and only if $\left(\lambda + \mu \frac{-f'(y-e^*)}{f(y-e^*)}\right)u_2 + \pi_E(y)(w_F + \alpha^+_{\psi,\rho_E,\rho_F}) \geq 0$.

We have that $y^*_{(\phi,\rho_E,\rho_F)} \mapsto +\infty$ when $\rho_F \mapsto 0$. In particular, there exists a threshold on ρ_F below which $\left(\lambda + \mu \frac{-f'(y^*_{(\phi,\rho_E,\rho_F)}-e^*)}{f(y^*_{(\phi,\rho_E,\rho_F)}-e^*)}\right)u_2 + \pi_E(y^*_{(\phi,\rho_E,\rho_F)})(w_F + \alpha^+_{\psi,\rho_E,\rho_F}) > 0$. In all these cases, the firm sends positive recommendation to employer ρ_E with strictly positive probability.

□

Overall, when the firm can re-hire the agent in the second period, it has an incentive to disclose information to the player who values it the most if it is itself or when it has sufficient bargaining power to extract the other employer's rent. However, in the other cases, there exist different forces with opposite effects on the firm's incentives.

First, if the firm has a bargaining power strictly below 1 the firm may choose not to disclose information and retain the agent, even if from employers value the information more than itself. The reason is that the firm would find it unprofitable to let another employer hire the agent because it could not extract enough rent.

On the contrary, even when the bargaining power is null, the firm may decide to recommend the agent to the other employers because it provides effort incentives in the first period.

1.7 Extensions and Discussion

1.7.1 Laissez-faire

In this subsection, we study the relationship between competition and reputation protection and show that disclosing information in contexts unrelated to the context where the information was collected is a sign of market power.

Consider a decentralized economy where, in the first period, two firms compete to attract the agent. There are no transfers and therefore the firms compete in disclosure policies. Furthermore, the firms cannot resale information to the employers downstream.

Specifically, assume that the firms are located at the extremities of a Hotelling line of size 1 with transportation cost s. Furthermore, as in Bénabou-Tirole (2016), that the agent's outside options, U_0, are also located at the extremities of the line so that she has to pay the transportation cost for her outside option as well. This assumption allows us to compare different levels of competitiveness - by varying the transportation cost s - without affecting the relative attractiveness of the outside option. After firms have made their offers, the agent is randomly allocated to a position $x \in [0,1]$ according to the uniform distribution. The second period is unchanged.

In equilibrium, each firm designs a disclosure policy of the form characterized in Proposition 2 and the agent chooses the one providing the highest intertemporal utility. To each disclosure policy ψ^* corresponds a unique intertemporal payoff for the agent $\mathcal{U}$ and a unique effort e^*, and therefore the problem of firm k can be written as follows. Given firm l's offer, U_l, firm k's probability to attract the agent when it offers $U_k \geq U_0$ is $\frac{1}{2} + \frac{U_k - U_l}{2s}$. Therefore, firm k chooses U_k so as to maximize her expected profit and thus solves the program

$$\max_{U_k=\mathcal{U}(e_k)} e_k\left(\frac{1}{2}+\frac{U_k-U_l}{2s}\right)$$

We select the unique symmetric equilibrium of the game and assume that an interior solution exists. An interior solution satisfies the first-order condition: $\mathcal{U}'(e^*)e^*+s=0$. Deriving this expression with respect to s, we get that $\frac{de^*}{ds}\geq 0$. That is, the equilibrium effort provision increases with the monopsony power. Let $y^*(s)$ denote the uniform cutoff associated with effort $e^*(s)$. $\frac{de^*}{ds}\geq 0$ is equivalent to $\frac{dy^*}{ds}\geq 0$.

We conclude in the following Proposition 6.

Proposition 6. *(Laissez-faire)*
Firms with greater monopsony power can incentivize higher effort by disclosing more information and to more employers, in particular to employers offering jobs less similar to the one where the information was generated

1.7.2 Output-contingent monetary compensation

So far, we have assumed that there are no transfers and focused attention on reputational incentives. In practice, however, we often observe a combination of both reputational and financial incentives. For instance, individuals are incentivized not to commit crimes with fines but also with their inscription on a criminal record that is then accessible to certain the firms. Similarly, government use naming-and-shaming policies to incentivize the firms to adopt or avoid certain behaviors on top of financial incentives such as fines for tax fraud or taxes on carbon emissions. In this subsection, we show that agency costs - such as liquidity constraints or risk aversion - are a necessary condition to rationalize reputational incentives.

Because our focus in on incentives, we consider the set-up of private contracting with no resale. We assume that the agent's date-1 output is verifiable and thus the firm and the agent can contract on it. The firm can commit both to a disclosure policy ψ and to a linear wage contract of the form $\phi(y)=a+by$, where y is the agent's output. We assume that the agent has separable (expected) utility between the first and the second period payoffs and that there is no discounting. Given a disclosure policy ψ and a wage schedule ϕ, the agent's expected payoff upon providing effort e is $E_{y,\rho}[u_1(\phi(y))+\psi(y,\rho)u_2]-c(e)$, where u_1 is concave.

The firm's program. The firm designs the disclosure policy ψ and the wage schedule ϕ to maximize date-1 profit. The program writes

$$\max_{\phi,\psi,e^*} \; e^* - E[\phi(y)]$$

$$\text{s.t.} \begin{cases} E[u_1(\phi(y))] + E_{y,\rho}[\psi_\rho(y)u_2] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = E_{y,\rho}\Big[\Big(u_1(\phi(y)) + \psi_\rho(y)u_2\Big)\Big(\dfrac{-f'(y-e^*)}{f(y-e^*)}\Big)\Big] & (IC_A) \\ \Pi^+_\rho \geq 0 & (IC^+_\rho) \end{cases}$$

We use a Lagrange approach and derive the first-order condition to obtain the equation characterizing the efficient disclosure policy

$$\psi_\rho(y) = 1 \iff \left(\lambda + \mu\frac{(-f'(y-e^*))}{f(y-e^*)}\right)u_2 + \big(\underline{\pi} + \rho E[\theta|y]\big)\alpha^+_\rho \geq 0 \tag{1.7}$$

This equation has the same form as in the case without transfers. We say that the firm resorts to reputational incentives if and only if $\mu > 0$. Otherwise, if $\mu = 0$, we are in the benchmark case of Section 1.5.1 and the firm protects the agent's reputation vis-à-vis all employers.

Risk neutrality. Assume that the agent is risk neutral with respect to financial risk; ie. $u_1(\phi) = \phi$. Computation is straightforward and we find that the optimal effort solves $c'(e^*) = 1$, which corresponds to the static first-best level of effort, and that $\mu = 0$.

The fact that the firm implements the static first best level of effort is intuitive because transfers are allowed. The firm could have chosen not to disclose any information, which corresponds to the textbook static model of moral hazard with no agency costs.

The fact that $\mu = 0$ means that the efficient disclosure policy is the benchmark disclosure policy, ψ^B. That is, when there is no agency cost, it is optimal for the firm to protest the agent's reputation vis-à-vis all potential employers and to incentivize effort only through output-contingent monetary transfers. The intuition is that reputational incentives are less efficient than financial ones because employers are not always obedient while there are no financial frictions. Furthermore, increasing the agent's reputational payoff on the second period is costless to the firm and relaxes the agent's participation constraint. The firm can then extract surplus with the fixed fee.

However, compared with the static game, ψ^B provides 'natural' incentives for effort because the agent is recommended more often following high output than low output. As a result, it is not optimal to make agent residual claimant (financially) otherwise she exerts too much effort, which lowers welfare and thus the profit that the firm can extract. Therefore, the optimal wage schedule is of the form $\phi(y) = a + by$, with $b < 1$.[19]

We conclude in the following Proposition 7.

Proposition 7. *(Reputational incentives)*

Agency costs are a necessary conditions for reputational incentives. In the absence of agency cost, it is efficient to provide incentives through transfers and protect the agent's reputation vis-à-vis all potential employers.

Proof. (proof of Proposition 7)

Because the agent is risk neutral, her incentive compatibility constraint (IC_A) simplifies to $c'(e^*) = b + E_{y,\rho}\Big[\psi_\rho(y)u_2\Big(\dfrac{-f'(y-e^*)}{f(y-e^*)}\Big)\Big]$.

The firm's program is

$$\max_{\phi,\psi,e^*} e^* - (a + be^*)$$

$$\text{s.t.} \begin{cases} a + be^* + E_{y,\rho}[\psi_\rho(y)u_2] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = b + E_{y,\rho}\Big[\Big(a + by + \psi_\rho(y)u_2\Big)\Big(\dfrac{-f'(y-e^*)}{f(y-e^*)}\Big)\Big] & (IC_A) \\ \Pi^+_\rho \geq 0 & (IC^+_\rho) \end{cases}$$

The first-order conditions are

$$(a) \qquad \lambda = 1$$

$$(b) \qquad \mu = 0$$

$$(e) \qquad 1 - b + (b - c'(e^*))\lambda - c''(e^*)\mu = 0$$

First, we get that $\mu = 0$ from FOC (b).

Second, using FOCs (a) and (b) we can rewrite FOC (e) into $c'(e^*) = 1$. This equation corresponds to that of the static first best.

Finally, rewriting the agent's incentive compatibility constraint (IC_A), we get:
$b = 1 - E[\psi_\rho(y)\frac{-f'(y-e^*)}{f(y-e^*)}] < 1$ □

Risk aversion. We now consider the case where the agent is averse to financial risk and show that the firm substitutes financial incentives with reputational incentives.

To make our point, we simplify the problem. First, we assume that there is a single employer in the second period, hiring for job ρ known to all players ex ante. Further-

[19] This result is a special case of Gibbons and Murphy (1992) who shows in a dynamic model that, in the presence of career concerns, the optimal wage schedule optimizes the sum of reputational and financial incentives.

more, to simplify computation we assume that the agent has mean-variance preferences; $E[u_1(\phi)] = E[\phi] - rVar(\phi)$, where r is the agent's risk aversion. Finally, recall that the production technology is $y = \theta + e + \varepsilon$. We assume that θ and ε are both normally distributed. Thus $\theta + \varepsilon$ is normally distributed, with mean zero, and let σ^2 denote the variance $\theta + \varepsilon$.

We know that the disclosure policy is a cut-off rule. Let y^* denote the cut-off chosen by the firm. The firm's program is

$$\max_{\psi, e^*} \; e^* - (a + be^*)$$

$$\text{s.t.} \begin{cases} (a + be^*) - rb^2\sigma_z^2 + (1 - F(z_A))u_2 - c(e^*) = U_0 & (IR_A) \\ c'(e^*) = b + f(z_A)u_2 & (IC_A) \\ z^* \geq z_\rho^+(e^*) & (IC_\rho^+) \end{cases}$$

First, consider the agent's incentive constraint and fix the effort target e^*. We get that $\frac{\partial z^*}{\partial b} < 0$. That is, financial and reputational incentives are substitutes.

Second, substitute (IR_A) and (IC_A) into the objective function and fix the effort target e^*. We have that $\frac{\partial^2 \text{Obj. function}}{\partial r \partial z^*} = 2b\sigma^2 f'(z^*) > 0$. Therefore, from Topkis' theorem, we conclude that $\frac{\partial z^*}{\partial r} > 0$; that is, the firm relies more strongly on reputational incentives the greater risk aversion is.

We conclude in the following Proposition 8.

Proposition 8. *(Financial vs. reputational incentives)*
As agency costs increase, reputational incentives substitute for financial incentives.

1.7.3 Essential services

So far we have focused on cases where audiences have heterogeneous valuations for information. We now turn to a case where heterogeneity comes from the agent's valuations for access to certain services. In particular, losing access to essential services significantly reduces one's welfare - such as access to housing, water and electricity, equal opportunities to get an education and a good job, faire access to loans and health care, fair treatment in justice, etc. Therefore, policy makers often pass laws to guarantee access to essential services, or at least to limit ostracism and discrimination. For instance, in France, landlords and electricity retailers are not allowed to hold blacklists of fraudulent tenants and customers and the European Commission proposes targeted intervention for regulating artificial intelligence in essential services.[20]

[20] Essential services are referred to as "'high-risk' AI use cases" by the European Commission in its proposal for an Artificial Intelligence Act.

To ease the exposition, we keep the notation of the model. Assume that at the beginning of period 2, Nature randomly draws an employer whose job is valued $u_2 > 0$ for the agent according to some distribution G on $\mathbb{R}^+$. Assume that all employers value information equally, with $\rho = 1$. Therefore, if they are given the same information about the agent, they value hiring the agent equally.

Suppose that a principal designs a disclosure policy to maximize a weighted average of all players' welfare. The principal's program writes

$$\max_{\psi^*, e^*} \; w_1 e^* + w_A \big(E_{u_2,y}[\psi_{u_2}(y) u_2 | e^*] - c(e^*) \big) + w_2 E_{u_2,y}[\psi_{u_2}(y) \big(\underline{\pi} + E[\theta | e^*] \big)]$$

$$\text{s.t.} \begin{cases} E_{u_2,y}\big[\psi_{u_2}(y) u_2 | e^*\big] - c(e^*) \geq U_0 & (IR_A) \\ c'(e^*) = E_{u_2,y}\big[u_2 \psi_{u_2}(y) \frac{(-f'(y-e^*))}{f(y-e^*)} | e^*\big] & (IC_A) \\ E_{u_2,y}[\psi_{u_2}(y) \big(\underline{\pi} + E[\theta | e^*] \big)] \geq 0 & (IC^+_{u_2}), \forall u_2 \geq 0 \end{cases}$$

Let $e^* > 0$ denote the effort target. The program can be solved by a first-order approach, which is valid under Assumption 1, and a slight adaptation of Assumptions 4 and 5 to account for heterogeneity in $u_2 \geq 0$ rather $\rho \in [0, 1]$. The disclosure policy is a cutoff rule on the agent's output. If employer u_2's obedience constraint binds, the cutoff is $y^+_{u_2}(e^*)$. Otherwise, the cutoff $y^*_{u_2}$ for employer u_2 solves the first-order condition Equation 1.8, below. Overall, the cutoff for employer u_2 is $\max\{y^+_{u_2}(e^*), y^*_{u_2}\}$.

$$\big(w_A + \lambda + \mu \frac{(-f'(y-e^*))}{f(y-e^*)} \big) u_2 + \big(\underline{\pi} + \rho E[\theta | y, e^*] \big) w_2 = 0 \tag{1.8}$$

Taking the derivative of Equation 1.8 with respect to u_2 we find that the cutoff $y^*_{u_2}$ for job u_2 strictly decreases with u_2 if and only if $w_2 > 0$. This means that a necessary condition to protect the agent's reputation in essential services is that the principal internalize the employers' profits. The reasoning is the following.

Recall that the efficient way to incentivize effort is through a cutoff contingent on performance only; ie. a uniform cut-off. Now that the principal accounts for date-2 profits, he must increase some, if not all, cutoffs. The choice of the cutoffs to increase has no effect on date-2 profits but the agent prefers to increase ostracism for the jobs she values the least. However, protecting the agent's reputation in priority for high-u_2 jobs dampens incentives. To see this, consider Equation 1.8 and note that $\frac{\partial^2}{\partial u_2 \partial(-y)} \mu \frac{(-f'(y-e^*))}{f(y-e^*)} \big) u_2 < 0$. Therefore, the agent's reputation is protected less vis-à-vis these employers than if there was no moral hazard.

We summarize these findings in Proposition 9.

Proposition 9. *(Essential services)*
Assume that the agent has heterogeneous preferences for access to services and that this is the sole source of heterogeneity. Suppose that a principal designs a disclosure policy to maximize a weighted average of all players' welfare. A necessary condition for differentiated disclosure of information is that the principal internalizes date-2 profits. Differentiated disclosure dampens incentives, but less so when the agent's reputation is protected in priority in essential services.

1.7.4 Discussion about existing laws and future regulation

In this subsection we first discuss several existing laws and rationalize them as outcomes of an information design problem as captured by our model. Then, we discuss how our model can inform the debates to regulate the design of algorithms.

Employer's access to personal data. In several countries the law regulates employers' access to information about their employees and about job candidates, such as criminal records, credit scores, profile on social media, drug tests, etc. The spirit of these laws is to ban access to such information, except for employers hiring in sensitive jobs with high risks of corruption - such as in the police force, the justice system, or for accounting positions - and where workers are in contact with vulnerable people, such as in retirement houses or schools. The design of and access to criminal records in France is an emblematic example because there exists three different versions listing different crimes and accessible to different parties.

This situation can be studied with our model in the case where a social planner maximizes welfare. Depending on the context, in the first period the agent takes out loans at a bank and must pay them back (credit score) or interacts with various individuals and must not commit crimes (criminal records). In the former case talent is the credit-worthiness and effort is costly because, for instance, the agent must forego consumption to reimburse her loans. In the latter case, the agent must control her temper and behave pro-socially which may require sacrificing personal welfare for others. In the second period, there are multiple employers hiring for various jobs. If, statistically, the risk of corruption is higher among workers who are in debt or who have committed crimes then employers hiring for sensitive jobs are the ones most interested in workers' credit scores or criminal records. The model predicts that, because the social planner internalizes date-2 profits, differentiated information is disclosed and employers hiring for sensitive jobs receive more information.

There exists a literature studying the heterogeneous effects of laws protecting workers' privacy by regulating access to information which jeopardizes their reputation such as financial credit reports (Bartik and Nelson, 2019; Corbae and Glover, 2018; Cortes et al., 2020), criminal records (Agan and Starr, 2018; Finlay, 2008), or drug tests (Wozniak, 2015). The effect of these policies typically vary across individual characteristics (such as education, sex and age) but also along job characteristics because the information concealed is more useful for some jobs than others. In line with the predictions of our model, Cortes et al. (2020) find that restricting employers' access to the credit reports of job applicants can destroy jobs because adverse selection becomes too severe. Job destruction is stronger where adverse selection is the highest - such as the in counties with a large share of subprime residents - and for jobs for which the credit scores are more relevant information - such as occupations requiring less than education or those that involve routine tasks.

Blacklists. Blacklists are documents recording consumers who have committed fraud. For instance, telecom retailers have blacklists for customers who do not pay their bills and supermarkets for consumers who have committed shoplifting. In France the data protection authority recommends that firms within a sector of activity are allowed to share a blacklist to prevent fraud, but firms across sectors are not.[21] The regulation is justified by the risk that firms abuse their market power and that consumers are exclude from excluding certain services.

This situation can be studied with our model in the case of private contracting without data resale. The agent is a customer contracting with a retailer in a given market in the first period and with other retailers in any market in the second period. The output is the number of bills paid to the first retailer and the cost effort corresponds to the cost of the bills and any foregone consumption to ensure the agent has enough budget at the end of the month. If the agent has many unpaid bills (low output) then she is blacklisted (negative recommendation), otherwise she is not (positive recommendation). Not paying one's telephone's bills is likely a signal of credit constraints or poor morals and this information can be relevant in several markets. The model predicts that the date-1 retailer has an incentive to disclose widely its blacklist to encourage effort.

Towards regulation. Our model can also inform policy debates about the regulation of data processing within multi-sided platforms and the regulation of artificial intelligence (AI).

First, the examples above show that various parties have an incentive to share information between themselves and some regulations already exist. In these cases, if there are no markets for information or if there are high transaction costs to write formal agreements to share in-

[21] See CNIL(2003) available at https://www.vie-publique.fr/sites/default/files/rapport/pdf/034000689.pdf

formation, the parties who initially collect information have an incentive to disclose it widely, or even publicly, to incentivize effort. In our model, this corresponds to the case of private contracting with no information resale. Multi-sided platforms, on the other hand, share information internally and therefore they internalize the welfare of all players.[22] It could be, however, that the relative weights that the platform puts on players' welfare do not coincide with those that regulators or society would choose. In that context, our model is relevant because it sheds lights on the trade-offs that platforms face, which can guide policy makers.

Second, the European Commission proposes to regulate AI to ensure it is human-centric and, in particular, it raises concerns about the risk of ostracism or discrimination in contexts which are unrelated to the context in which the data was generated. Our model is relevant for address these questions because we study how the interest of the agent are traded-off with those of other players she interacts with and we explicitly allow for heterogeneity, which captures the idea that the contexts where information is collected and disclosed may be more or less related. The European Commission gives special attention to 'high-risk' situations such as for recruitment, judicial decision making or the evaluation of credit worthiness, which our model accounts for in the extension for essential services. Finally, another concern raised by the EC is that algorithms can be opaque and thus regulators do not know how the information used was obtained, even if they could observe it. In that regard, our model is relevant because it analyzes how incentives affect the design of some of these algorithms.

1.8 Conclusion

In this paper we study the disclosure of information about an agent's output to audiences with heterogeneous valuations for such information and ask when it is efficient to differentiate information disclosure. In our model, disclosing information to any audience provides incentives ex ante and improves screening ex post but hurts the agent because effort is costly and she is excluded more often. The policy maximizing effort discloses as much information as possible and as widely as possible, subject to the agent's participation constraint. As a result, differentiated disclosure dampens incentives. We characterize the optimal disclosure policy for different objectives and we identify three rationales for differentiated information disclosure.

First, in order to protect the agent's reputation, information must be tailored to the audi-

[22]For instance, Facebook to merge all the databases of WhatsApp, Messenger and Instagram and, in Germany, the regulator banned Facebook from processing WhatsApp user data for other Facebook services. The case was covered, for instance, by https://www.reuters.com/business/legal/german-regulator-bans-facebook-processing-whatsapp-user-data-2021-05-11/

ences because they use information differently. Second, if the audiences' payoffs are internalized then it is efficient to provide more information to those who value it the most. Finally, if the agent has heterogeneous preferences for audiences then it is efficient to provide more information to those who value it the most but only if the audiences' payoffs are internalized.

Our model can rationalize existing laws protecting workers and consumers privacy as outcomes of an information design problem. It can also be used to address questions related to the regulation of algorithms and the processing of information by multi-sided platforms because it highlights the effect of various incentives on the design of some of these algorithms.

1.A Proofs

1.A.1 Proof of Proposition 1

We prove that the first-order approach used to characterize efficient disclosure policies in the baseline model with a single audience and no information resale in Proposition 1 is valid. For that, we must show that e^* is a global maximum of the agent's program.

We have shown that, if the FOA is valid, a disclosure policy is a cutoff rule on output. Let (y^*, e^*) denote the cutoff and associated effort. We now show that, under our assumptions, when the disclosure policy is a cutoff rule as defined above, it is optimal for the agent to choose effort e^*. The agent takes the cutoff y^* as given and chooses her effort provision e so as to maximize her intertemporal payoff. We have

$$\begin{aligned} U(e) &= \Big(1 - F(y^* - e)\Big)u_2 - c(e) \\ U'(e) &= f(y^* - e)u_2 - c'(e) \\ U''(e) &= \big(-f'(y^* - e)\big)u_2 - c''(e) \end{aligned}$$

By assumption, f is singled-peaked at 0 and therefore $\big(-f'(y^* - e)\big)$ has same sign as $y^* - e$. We have shown that, in equilibrium, the firm chooses $y^* - e^* \leq 0$ and therefore, $U''(e^*) < 0$. This proves that e^* is a local maximum. Furthermore, for all $e \geq e^*$, $y^* - e < y^* - e^* \leq 0$ and thus $U''(e) < 0$. Therefore, we need only find conditions such that no $e < e^*$ can be a profitable deviation.

• Strong assumption. Condition such that the agent's program is concave: $\forall e < \bar{e}$, $U''(e) < 0$:

$$\min_{0 \leq e \leq \bar{e}} c''(e) \geq \max_{z \in \mathbb{R}} \big(-f'(z)\big)u_2 \tag{9}$$

This is a strong assumption which guarantees that the cost of effort is so convex that the agent's utility is necessarily concave in effort.

• Mild assumption. Conditions such that the agent's program is single-peaked: $\forall e < e^*$, $U'(e) \geq 0$.

First, if $y^* - e \in [y^* - e^*, e^* - y^*]$

$$\begin{aligned} U'(e) &= f(y^* - e)u_2 - c'(e) \\ &> f(y^* - e^*)u_2 - c'(e) \\ &= c'(e^*) - c'(e) > 0 \end{aligned}$$

See Figure 5. Therefore, if $y^* - e \in [y^* - e^*, e^* - y^*]$ for all $e < e^*$ then the agent's program, for y^* fixed, is singled peaked at e^*.

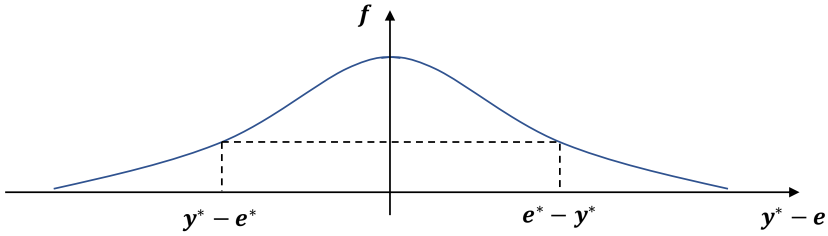


Figure 5 – Distribution of f.
Note. For $e < e^*$ we have that $y^* - e > y^* - e^*$. If $y^* - e \in (y^* - e^*, e^* - y^*)$ then $f(y^* - e) > f(y^* - e^*) = c'(e^*)$. Otherwise, if $y^* - e \geq e^* - y^*$ then we must find another lower bound for $f(y^* - e)$ and account for the fact that f is a decreasing function on $\mathbb{R}^+$.

However, it could be that there exists $e > 0$ and y^* implementable such that $y^* - e > e^* - y^*$. This is the case, for instance, for $y^* = \overline{y}$, the cut-off inducing the highest effort. [23] Note that if $y^* - e > e^* - y^*$ then $y^* - e > 0$. Given that f is single-peaked at 0, we have that $f(y^* - e)u_2 - c'(e) > f(\overline{y} - e)u_2 - c'(e)$. Therefore, it is sufficient that $f(\overline{y} - e)u_2 - c'(e) \geq 0$ for all $e < 2y^* - e^*$. This condition must be satisfied for all candidate (y^*, e^*) with $y^* - e > e^* - y^*$ and the most demanding case is for $(\overline{y}, \overline{e})$. Therefore the assumption writes:

$$f(\overline{y} - e)u_2 - c'(e) \geq 0, \forall e \in [0, \overline{e}]$$

[23] For instance, assume that the agent's participation constraint does not bind when the cut-off induces the highest effort $\overline{y} = \overline{e} > 0$. Then $\overline{y} - \epsilon > 0 = \overline{y} - \overline{e}$ for ϵ small enough.

Chapter 2

Match me if you can

Abstract

A principal maximizing welfare on a platform can design matching mechanisms to screen types and incentivize effort. On the platform, high-type agents are a valuable and scarce resource because they provide a higher payoff in a match. In a two-period model, we show that the principal optimally offers a separating menu of contracts that differ both in quantity and quality of matches. Low-type agents choose the contract offering few matches but with high-type agents only while high-type agents choose the contract offering the most matches. A platform entering the market in date 2 attracts high-type agents by screening low-type agents more than the incumbent does when it is a monopolist. Platform entry thus creates a hold-up problem which reduces effort provision because an entrant does not internalize its effect on pre-entry incentives. This issue is exacerbated with data portability because the entrant can screen more efficiently which further reduces incentives for effort.

2.1 Introduction

In this paper, I explore the idea that a platform can design its algorithms not only to help individuals find each other but also to solve information failures. It rests upon the assumption that some individuals are better match partners than others - for instance because they behave more altruistically - and this paper asks who to match them with.

In the model, a principal designs a matching mechanism between two symmetric sides, each composed of a mass one of agents. There are no transfers. Agents can accept to interact with the partner the platform has matched them with or they can choose their outside option. Upon interacting, each agent chooses an effort level $e \in \{0, 1\}$, which has a positive externality 1 on their partner. There are two types of agents, high and low. We assume that high-type agents benefit more from using the platform than low-type individuals and always contribute $e = 1$.[1] On the contrary, low-type agents incur a cost for effort and therefore must be incentivized to contribute $e = 1$. We study the optimal matching mechanism of a principal maximizing welfare on its platform when types are private information in the static case, in a two-period model with and without entry, and with and without data portability. The main take-aways are the following.

As a benchmark, we consider a static game with symmetric information. Then, the principal implements positive assortative matching (PAM): agents are matched with others of the same type, high types stay on the platform and low types get their outside option. However, in the static game, when agents are privately informed about their type, the principal screens low-type agents through contracts with heterogeneous match quantity and match quality. High-type agents are always accepted on the platform but may be matched with low-type agents. Low-type agents are not always accepted on the platform but they are only matched with high-type agents such that, at the optimum, their incentive compatibility constraint binds. That is, the principal implements the matching closer to PAM, conditional on leaving just enough rent to low-type agents to learn their type.

In the two-period case, it is optimal to accept all agents on the platform in the first period and incentivize them to contribute $e = 1$, and exclude as many low-type agents as possible in the second period. There are two regimes to consider, depending on whether the share of high-type agents in the population is high - and then adverse selection dominates - or mild - and then moral hazard dominates. When the share of high-type agents is high, the principal can implement the optimal static mechanism in date 2 and random matching in date 1. Date-2 payoffs on the equilibrium path are sufficiently high to discourage agents from shirking

[1] Equivalently, one can assume that high-type agents are altruistic and partially internalize the partner's payoff.

because they would then be excluded in date 2. We say that in this regime adverse selection dominates because the principal only needs to deviate from PAM to screen types as in the static case and this is sufficient to incentivize effort. When the share of high-type agents is mild, the optimal static mechanism does not provide sufficient date-2 payoffs to incentivize effort in date 1. Then, the principal must accept more low-type agents in date 2 and, to minimize their number, matches them exclusively with high-type agents. We say that in this regime moral hazard dominates because, to incentivize effort, the principal needs to deviate from PAM more than just to screen types.

Finally, we extend the dynamic model by allowing for platform entry in the second period. There exist multiple equilibria so we focus on entry-proof mechanisms. A mechanism designed by the incumbent is said to be entry-proof if there exists an equilibrium in which all high-type agents remain on the incumbent's platform in date 2 and obtain the highest equilibrium payoff. In a first step, we consider the case without data portability: the entrant knows the incumbent's mechanism but not agents' date-1 behavior. The entrant designs a matching mechanism to maximize welfare on its platform and therefore designing the optimal static mechanism is a dominant strategy.

In the regime where adverse selection dominates, the optimal dynamic mechanism with no entry and the entrant's mechanism are identical in period 2 and therefore the optimal dynamic mechanism is entry-proof. However, in the regime where moral hazard dominates, the optimal dynamic mechanism with no entry leaves rent to low-type agents compared with the entrant's mechanism. The entrant's mechanism provides a strictly higher payoff to high-type agents and thus the optimal dynamic mechanism with no entry is not entry-proof. That is, entry creates a hold-up problem because low-type agents are no longer compensated for their date-1 effort. We fully characterize the optimal dynamic mechanism with entry and no data portability and show that the incumbent must allow some low-type agents to shirk in date 1 and exclude in date 2 so as to provide as much date-2 payoff to high-type agents as with the entrant's mechanism.

We then allow agents to port their ratings to prove to the entrant whether or not they contributed $e = 1$ in date 1, which allows the entrant to better screen types because only low-type agents do not contribute in date 1. Now, the dominant strategy is to design the optimal static mechanism but restricted to the set of agents who have contributed in date 1. In the regime where adverse selection dominates, the optimal dynamic mechanism with no entry is still entry-proof. However, in the regime where moral hazard dominates, the optimal dynamic mechanism without data portability is no longer entry-proof because it still left some rent to low-type agents. We fully characterize the optimal dynamic mechanism with entry and data portability and show that the incumbent must allow even more low-type agents to shirk in date 1 than in the case without data portability.

2.1.1 Literature Review

Related Literature. Our paper contributes to the literature on non-monetary incentives by combining elements from matching theory and mechanism design. The key innovation is to introduce future matches as the valuable outcome motivating agents.

As in Holmström (1999)'s model of career concerns, the agent exerts effort today to improve its match opportunities tomorrow. However, in our model incentives are explicit in the sense that the principal controls future matches while in Holmström (1999) the agent's output is publicly observable and therefore incentives are implicit. More generally, our paper relates to a large and growing literature in information design which studies how the disclosure of information ex post changes incentives ex ante. For instance, Harbaugh and Rasmusen (2018) study the optimal design of grades by certifiers of quality when firms can choose whether to apply for certification. The closest paper to ours is Rodina (2016) who builds a two-period model with an agent with career concerns and allows the principal to control which information to disclose to the market in date 2. Thus, in their paper, the principal is constrained by Bayes' rule while in ours the key constraint is the matching constraint. In the former, the principal cannot systematically fool the market into believing that the agent performed well and must therefore strategically choose what to disclose the market. In the latter, the scarce resource is the quantity of high-type agents and the principal must strategically choose who to match them with.

Our paper also relates to the literature on information sharing among competitors. Mukherjee (2008) study information disclosure in a dynamic model of career concerns where an agent works for an incumbent but where raiders can try to attract the agent but where the incumbent is better informed about the agent's abilities. Padilla and Pagano (1997) show that competing lenders can benefit from sharing information about borrowers to discipline them. In our paper we study how the threat of platform entry affects the design of the matching mechanism and compare the cases with and without data portability.

Finally, our paper borrows from the large literature about matching and surveyed for instance in Chade et al. (2017). In particular, a part of literature studies how the prospect of future matches affects incentives to invest ex ante. For instance, Nöldeke and Samuelson (2015) finds condition for positive assortative matching to generate efficient incentives. Contrary to this literature where matches are decentralized, in our paper matches are perfectly controlled by a principal who can use them to incentivize effort.

2.2 Static model

2.2.1 The model

Consider an economy with two sets of agents, each of mass 1. A Principal designs a mechanism ϕ to match agents by pairs, denoted (i, j).

Upon being matched with j, agent i either accepts or refuses the interaction. In each pair, if at least one agent refuses to interact then both agents get their outside option $\underline{U} \geq 0$, independent of their type. If both agents accept the match then they interact. Upon interacting with agent j, agent i gets utility v_i and chooses an effort level $a_i\{0, 1\}$. Contributing $a_i = 1$ costs her $c_i \geq 0$ but generates a positive externality normalized to 1 on her partner. Overall, in each period, agent i of type θ_i interacting with agent j gets flow utility $u_i - c_i a_i + a_j$.

We assume that there are two types of agent in the economy, L and H, and there is a proportion σ of type-H agents. An agent's type refers to (u_i, c_i), where u_i is the payoff from interacting with another agent and c_i is the cost of effort. We assume that $(u_L = 0, c_L = c)$ where $c > 0$ and $(u_H = u, c_H = 0)$ where $u > 0$. In particular this means that high-type agents don't need to be incentivized to contribute $a = 1$.[2]

Matching Mechanism.

In the static game there are no incentive contract and therefore the principal cannot incentivize the low-type agents to exert effort. However, he can screen agents on the platform. At the beginning of the game, the principal commits to a matching mechanism ϕ. Building from Myerson (1982) we can, without loss of generality, restrict our analysis to incentive compatible direct mechanisms. Thus, at the beginning of period 1, the Principal asks agents to report their types. An agent reporting type $i \in \{L, H\}$ is matched with a (reported) low type with probability λ_i and with a (reported) high type with probability η_i.

To make the problem interesting we make the following assumptions.

Assumption 6. *(i)* $\underline{U} < 1$
(ii) $\underline{U} < u$

[2]Alternatively, we could assume that $u_H = 0$ and $c_H < 0$. This can be micro-founded, for instance, by assuming that high-type agents are altruistic and (partially) internalize their partner's payoff. With this specification, altruism leads high-type agents to always contribute and they also obtain higher payoff from being matched.

The first assumption implies that all agents prefer being matched to a high-type agent rather than getting their outside option. The second assumption implies that high-type agents prefer being accepted on the platform, regardless of who they are matched with, than getting their outside option. Together with the assumption that they always contribute $a = 1$, this implies that the principal needs only focus on low-type agents (for screening).

Matching constraint. The following accountability rule, called the matching constraint, must be satisfied: the number of agents matched with high-type agents is equal to the number of high-type agents accepted on the platform; $(\lambda_H + \eta_H)\sigma = \sigma\eta_H + (1-\sigma)\eta_L$. This equation simplifies to

$$(MC) \qquad\qquad \lambda_H \sigma = (1-\sigma)\eta_L$$

2.2.2 Feasible allocations

Assume that there is no asymmetric information and that each agent's type is public knowledge. We characterize the set of implementable allocations and the Pareto frontier.

Participation constraint.
Let $V_i = \lambda_i u_i + \eta_i(u_i + 1) + (1 - \lambda_i - \eta_i)\underline{U}$ denote the payoff of an agent with type $i \in \{L, H\}$. Agents can always choose their outside option and therefore it cannot be that $V_k < \underline{U}$. Thus the principal's matching mechanism must satisfy the following two participation constraints

$$V_L \geq \underline{U} \qquad\qquad (IR_L)$$

$$V_H \geq \underline{U} \qquad\qquad (IR_H)$$

Feasible allocations. Excluding all agents from the platform yields the lowest implementable payoff for all agents, $(\underline{U}, \underline{U})$, represented by point A on Figure 6.

The allocation that maximizes the high-type agents' payoff matches them together and excludes all the low-type agents from the platform, which yields $(u + 1, \underline{U})$, represented by point B on Figure 6. Any allocation on the line connecting A to B is feasible by accepting more and more high-type agents but no low-type agents. And no allocation below this line are feasible because low-type agents can always obtain $\underline{U}$ by choosing their outside option.

Low-type agents prefer their outside option rather being matched with another low-type agent. Therefore, the allocation that maximizes low-type agents' payoff matches them with high-type agents only, and excludes low-type agents if there are more than there are high type ones. This yields (u, V_L^{max}) where $V_L^{max} = 1$ if $\sigma \geq 1/2$ and $V_L^{max} = \frac{\sigma}{1-\sigma} + \big(1 - \frac{\sigma}{1-\sigma}\big)\underline{U}$, otherwise. This allocation is represented by point C on Figure 6. Any allocation on the line connecting A to C is feasible by accepting more and more high-type agents and matching them with low-type agents only and excluding the remaining low-type agents.

Finally, high-type agents prefer to participate on the platform rather than any other alternative and all agents prefer to match with them rather than getting their outside option. Therefore, on the Pareto frontier, all high-type agents are accepted on the platform; ie. $\lambda_H + \eta_H = 1$. On the contrary, low-type agents prefer their outside option rather being matched with another low-type agent and therefore, on the Pareto frontier, low-type agents are never matched together; ie. $\lambda_L = 0$. These two findings, combined with the matching constraint $\lambda_H \sigma = (1-\sigma)\eta_L$, imply that V_H and V_L are both linear functions of η_L and therefore the Pareto frontier in the static game is a line connecting B to C, obtained by increasing the share of low-type agents accepted on the platform; see Figure 6.

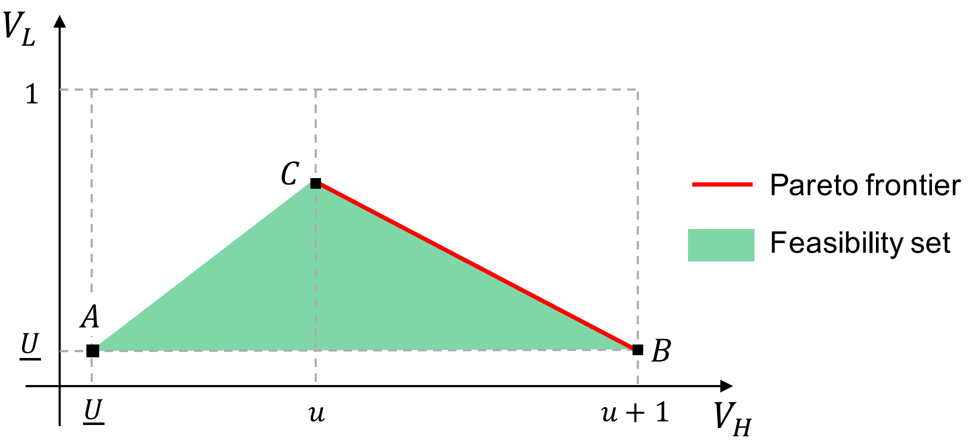


Figure 6 – Feasibility set and Pareto frontier - represented for $\sigma < 1/2$

Lemma 1. *(Static case - Symmetric information) In the static case, low-type agents are never matched together $\lambda_L = 0$ and all high-type agents are accepted on the platform $\lambda_H + \eta_H = 1$. Expected payoffs are $V_H = 1 + u - \frac{1-\sigma}{\sigma}\eta_L$ for high-type agents, and $V_L = (v+1)\eta_L + \underline{U}(1-\eta_L)$ for low-type agents, for all $\eta_L \in [0,1]$.*

Monopolist. Assume that the principal is a monopolist and designs a matching algorithm to maximize welfare on the platform. The principal's program writes

$$\max_{\eta_H,\lambda_H,\eta_L,\lambda_L} \sigma\big[v\lambda_H + (v+1)\eta_H\big] + (1-\sigma)\eta_L$$

$$\text{s.t. } \lambda_H\sigma = (1-\sigma)\eta_L \qquad (MC)$$

$$V_L \geq \underline{U} \qquad (IR_L)$$

First, it must be that all high-type agents are accepted on the platform. Otherwise, the principal can accept more high-type agents and match them together (by increasing η_H) which strictly increases welfare on the platform. Using the fact that $\lambda_H + \eta_H = 1$ and substituting the matching constraint into the objective function, the principal's program writes

$$\max_{\eta_H,\lambda_H,\eta_L,\lambda_L} (v+1)\sigma$$

$$\text{s.t. } V_L \geq \underline{U} \qquad (IR_L)$$

There exists multiple solutions to these problem. To select an equilibrium, assume there is an arbitrary small transaction cost associated with accepting an agent on the platform. Then, there exists a unique solution to principal's program. The principal accepts only the high-type agents on the platform and rejects all the low-type agents. This allocation corresponds to point B on Figure 6.

This corresponds to what the matching literature refers to as positive assortative matching (PAM) because agents of the same type are together - and, in our case, the low-type agents refuse the match. We conclude in Lemma 2.

Lemma 2. *(Static case - Monopolist) In the static case under symmetric information, a monopolist maximizing welfare on the platform implements positive assortative matching (PAM). That is, agents are matched with agents of the same type and the low-type agents refuse their match and choose their outside option instead.*

2.2.3 Asymmetric information

Assume now that agents privately know their types. We have shown that in a Pareto-efficient allocation low-type agents are never matched together. This is only feasible if pooling is not an equilibrium and therefore we will first consider separating allocations.

Incentive compatibility constraints. It is in the agents' best interest to report their types truthfully if and only if the following two incentive compatibility constraints are satisfied

$$V_L \geq V_H - (\lambda_H + \eta_H)u \qquad (IC_L)$$

$$V_H \geq V_L + (\lambda_L + \eta_L)u \qquad (IC_H)$$

We plot the set of constrained-efficient allocations, belonging to the Pareto frontier characterized in Lemma 1 and satisfying the incentive compatible allocations on Figure 7.

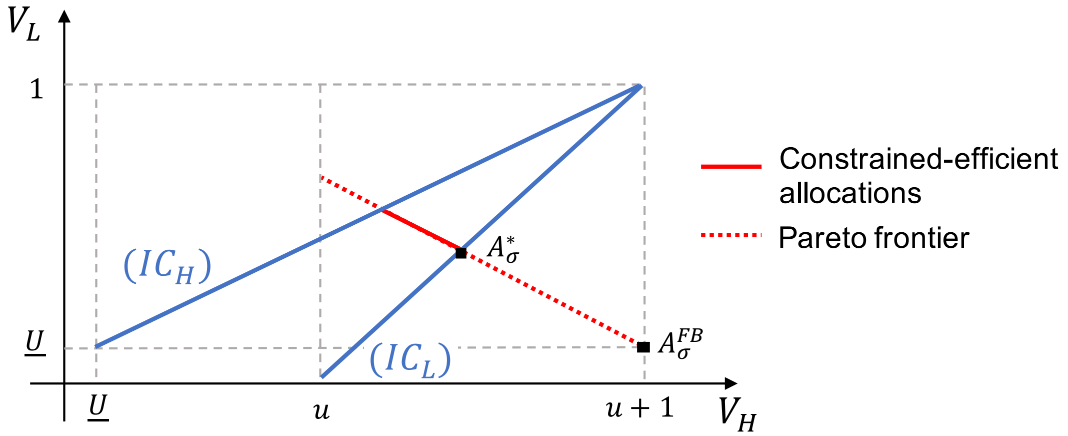


Figure 7 – Constrained-efficient allocations when types are private information (red full line). The red dotted line is the Pareto frontier in the unconstrained problem.

Note. Let σ denote the share of high-type agents in the population. A_σ^{FB} is the first best allocation and A_σ^* is the second best allocation.

For (IC_L) to be satisfied, it is necessary that the principal keeps sufficiently many low-type agents on the platform. Otherwise, there are so few low-type agents that it is profitable to pretend to be a high-type agent.

For (IC_H) to be satisfied, it is necessary that the principal excludes sufficiently many low-type agents. The reason is that the principal guarantees the low-type agents that they will only be matched with high-type agents, which provides higher payoff.

Lemma 3. *(Static case - Asymmetric information) In the static game when types are private information, pooling is inefficient. Furthermore, under a constrained-efficient matching mechanism, agent's expected payoffs are* $V_H = 1 + u - \frac{1-\sigma}{\sigma}\eta_L$ *for high-type agents, and* $V_L = (v+1)\eta_L + \underline{U}(1-\eta_L)$ *for low-type agents, for all* $\eta_L \in [0,1]$ *satisfying* (IC_L) *and* (IC_H).

Second best. Assume that the a monopolist designs a matching algorithm to maximize welfare on the platform. The monopolist's program writes

$$\max_{\eta_H,\lambda_H,\eta_L,\lambda_L} \sigma\big[v\lambda_H + (v+1)\eta_H\big] + (1-\sigma)\eta_L$$

$$\begin{aligned}
\text{s.t. } & V_L \geq \underline{U} && (IR_L)\\
& V_H \geq \underline{U} && (IR_H)\\
& V_L \geq V_H - u && (IC_L)\\
& V_H \geq V_L + \eta_L u && (IC_H)\\
& \lambda_H\sigma = (1-\sigma)\eta_L && (MC)
\end{aligned}$$

Let ϕ_σ^* denote the optimal static matching mechanism. Contrary to the case with symmetric information, PAM is no longer optimal because it violates the low-type agent's incentive compatibility constraint, (IC_L). That is, the monopolist must keep some low-type agents to encourage them to reveal their types truthfully. However, ϕ_σ^* is the closest Pareto-efficient static mechanism to PAM in the sense that it maximizes the probability to match high-type agents together.

Proposition 10. *(Static case - Monopolist) In the static case, when types are private information, positive assortative matching (PAM) does not satisfy the incentive compatibility constraint of low-type agents. To separate types and minimize the number of low-type agents on the platform, a monopolist designs two types of contracts which differ in the quantity and the quality of the matches.*

High-type agents are always accepted on the platform but are matched with a low-type agent with high probability. Low-type agents are not always accepted on the platform but are matched only with high-type agents. At the optimum, low-type agents are indifferent between both contracts.

Implementation. A practical example of separating contracts with a trade-off quantity vs. quality of matches is found on Airbnb which is a platform matching guests who want to rent an accommodation with hosts who can rent theirs. Hosts can use either 'instant book' or 'manual book'. Under 'manual book' hosts can screen guests by checking information and discussing with them prior to accepting their offers, but they are ranked lower than those opting for 'instance book', which means that they are less likely to be matched. On the contrary, under 'instant book' any request is automatically accepted but hosts are more likely to be matched.

2.3 Two-period model

2.3.1 Model

In the static game, the principal cannot incentivize effort because it cannot write incentive contracts. We now consider a two-period model (with no discounting) in which the static game studied in the previous section is repeated, which allows the principal to condition date-2 matches on date-1 outcomes. The timing of the game is the following.

At the beginning of period 1, the principal commits to a matching mechanism ϕ. Agents are asked to reveal their types to the platform and they send a message $i \in \{L, H\}$ to the platform. Based on their message, the platform forms matches according to ϕ and the static game studied in the previous section occurs. Agents either accept or reject the match. If at least one agent in a pair refuses the match then both agents are excluded from the platform for the rest of the game and they get their outside option, which provides payoff $\underline{U}$ in each period. If both agents in a pair accept their match then they interact. Upon interacting, agents choose an effort level $e \in \{0, 1\}$; contributing $e = 1$ has a positive externality valued 1 by the partner. At the end of period 1, agents report whether they received the externality or not; by assumption they report truthfully.

In the second period, the principal receives the reports and learns which agents contributed $e = 1$. Based on their date-1 effort provision and their date-1 message, the principal forms matches based on ϕ. As in the previous period, agents accept whether to interact or not and, upon interacting, choose their effort provision.

2.3.2 Analysis

In this paper, we will restrict attention to parameter values such that, in the first period, the principal can accept all agents to join the platform and contribute $e = 1$. As we will show later in greater details these situations occur when there is a sufficiently large share of high-type agents in the population. To ease the exposition, and in particular the number of variables in the program, we will now assume two results that can be proven with straightforward ex post computation.

First, if it is feasible for the principal to accept all agents to join the platform and contribute $e = 1$ then he strictly prefers to do so. The intuition for this result is that, in the worst case scenarios, when the principal needs to leave rent to low-type agents to incentivize effort, they are indifferent between contributing $e = 1$ in date 1 and have a chance to be on the platform in date 2 and not contributing in date 1 and be rejected in date 2. While they are indifferent, contributing $e = 1$ has a positive externality in date 1 and accepting them on the platform has no welfare effect in date 2 because we do not model transaction costs explicitly. Therefore we will assume up front that the principal wants to accept all agents to join the platform and contribute $e = 1$ in the first period.

As a result, the matching mechanism applied in date 1 is irrelevant because any date-1 matching leads to the same outcome and we need only characterize the date-2 mechanism. To ease notation, we will use the same notation as in the static case without specifying that it applies only to date 2; ie. $(\lambda_L, \eta_L, \lambda_H, \eta_H)$ for the mechanism and V_L and V_H for the payoffs.

Second, a similar reasoning shows that a constrained-efficient date-2 allocation belongs to the Pareto frontier and therefore that $\lambda_L = 0$ and $\lambda_H + \eta_H = 1$. From now on we will assume that this is the case. We now list the constraints involved in the principal's program.

Incentive compatibility constraints. Agents must reveal their types truthfully at the beginning of the first period. Note that, if all agents are on the platform in date 1 and all contribute $e = 1$ then agents' date-1 payoffs are independent of the type they have reported. Thus, low-type agents truthfully reveal their types in the period if and only if $1 - c + V_L \geq 1 - c + V_H - u$. This simplifies to $V_L \geq V_H - u$, which is exactly the incentive compatibility constraint for low-type agents in the static game studied earlier. Similarly, high-type agents reveal their types if and only if $V_H \geq V_L + u$. We keep the same notation as in the static game, (IC_L) and (IC_H).

Date-1 effort provision. In period 1, it must be in low-type agent's best interest to contribute $e = 1$. The harsher punishment, if they don't, is to reject them from the platform in period 2. The constraint writes $1 - c + V_L \geq 1 + \underline{U}$, which simplifies to $V_L \geq c + \underline{U}$. Let (IC'_L) denote this constraint.

Participation constraints. By assumption, high-type agents prefer to participate on the platform regardless of who they are matched with. Therefore, high-type agents' participation constraint is automatically satisfied in equilibrium. In period 2, the principal wants to reject as many low-type agents as possible and therefore there is no date-2 participation constraint. In period 1, the principal wants to accept all the low-type agents, therefore it is necessary that their expected payoff on the platform is greater than getting their outside option for both periods; $1 - c + V_L \geq 2\underline{U}$. Given (IC'_L) and Assumption 6-(i) we have $1 - c + V_L \geq 1 - c + c + \underline{U} = 1 + \underline{U} > 2\underline{U}$. Therefore, low-type agents' participation constraint is automatically satisfied in equilibrium.

The Principal's program. The principal designs a matching algorithm to maximize intertemporal welfare on the platform. The program writes

$$\max_{\eta_H,\lambda_H,\eta_L,\lambda_L} \sigma\big[(v+1)+v\lambda_H+(v+1)\eta_H\big]+(1-\sigma)(1+\eta_L)$$

$$\begin{aligned}
\text{s.t. } & V_L \geq V_H - u && (IC_L)\\
& V_H \geq V_L + \eta_L u && (IC_H)\\
& V_L \geq c + \underline{U} && (IC'_L)\\
& \lambda_H \sigma = (1-\sigma)\eta_L && (MC)
\end{aligned}$$

The computation to solve this program is straightforward and the solution is better seen graphically. The constraints (IC_L), (IC_H) and (IC'_L) are independent of σ. On Figure 8 we represent the three constraints as well as the date-2 optimal allocations for all $\sigma \in [\bar{\sigma}, \bar{\bar{\sigma}}]$. There are two regimes and either adverse selection dominates ($\sigma \geq \bar{\bar{\sigma}}$) or moral hazard dominates ($\sigma \in [\bar{\sigma}, \bar{\bar{\sigma}}]$).

When the share of high-type agents is sufficiently high ($\sigma \geq \bar{\bar{\sigma}}$) then applying the optimal static mechanism in the second period provides sufficient incentive for effort in the first period and is therefore optimal in the dynamic game as well. Formally, if (IC_L) is satisfied then (IC'_L) is slack and we say that adverse selection dominates; see Figure 8.

This is no longer the case for intermediate shares of high-type agents in the population ($\sigma \in [\bar{\sigma}, \bar{\bar{\sigma}}]$). Indeed, if (IC_L) binds then (IC'_L) is violated and therefore applying ϕ^*_σ in period 2 provides insufficient incentives for effort in period 1. The platform must provide then more rent to low-type agents in period 2 than would be optimal in the static case.

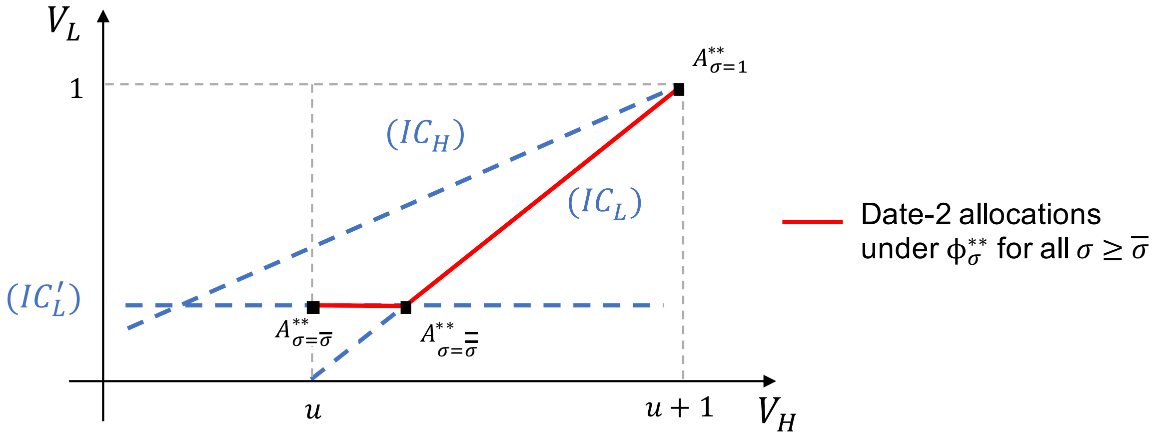


Figure 8 – Second-best date-2 allocations in the two-period game. If $\sigma \geq \bar{\bar{\sigma}}$ then adverse selection dominates and $A^{**}_{\sigma} = A^{*}_{\sigma}$. Otherwise, if $\sigma \in [\bar{\sigma}, \bar{\bar{\sigma}}]$ then moral hazard dominates and $A^{**}_{\sigma} \neq A^{*}_{\sigma}$.

Note. On this figure, (IC_H) and (IC'_L) cross when $V_H < u$; this is the case if and only if $\frac{\underline{U}+(1-\underline{U})u}{2-\underline{U}} \geq c + \underline{U}$. Then, $\bar{\sigma}$ is such that the Pareto frontier - which depends on σ, (IC'_L) and $V_H = u$ intersect at the same point. Otherwise, $\bar{\sigma}$ is such that the Pareto frontier - which depends on σ, (IC'_L) and (IC_H) intersect at the same point.

Discussion. In the static game studied previously, we have shown that PAM was not optimal for a monopolist because it violates low-type agents' incentive compatibility constraint. In order to screen types, a monopolist must design matching contracts that deviate from PAM. We have shown that implementing the static mechanism in the second period of the two-period game is also optimal when there are sufficiently many high types ($\sigma \geq \bar{\bar{\sigma}}$) in the population because then low-type agents have sufficiently high chances to be matched with one if they contribute $e = 1$ in the first period. However, when there are fewer high-type agents ($\sigma \in [\bar{\sigma}, \bar{\bar{\sigma}}]$) then, in order to incentivize effort, the monopolist must deviate further from PAM and increase the probability to match low-type agents with high-type agents.

Remark. Up to now, we have restricted attention to the case where the share of high-type agents in the population exceeds some threshold; $\sigma \geq \bar{\sigma}$. When it is below the threshold ($\sigma < \bar{\sigma}$) then the principal must allow some low-type agents to shirk in date 1. The mechanism is such that low-type agents are indifferent between contributing $e = 1$ in date 1 and then having a chance of being matched with a high-type agent in date 2, or shirking in date 1 and then being rejected from the platform in date 2. In all these cases, the share of high-type agents on the platform in date 2 is equal to $\bar{\sigma}$.

2.4 Entry in the second period

The previous sections have shown that a monopolist platform implements matching mechanisms that deviate from PAM - the optimal static mechanism under symmetric information. This implies that a monopolist must reduce the payoffs of high-type agents to better screen in date 2 and incentivize effort in date 1. While implementing these mechanisms is optimal from the perspective of the monopolist, matching high-type agents with low-type agents is inefficient ex post. Therefore in this section we allow for period-2 entry.

A platform may enter the economy in period 2 and attract agents who were on the incumbent's platform in period 1. We will compare situations without and with data portability; that is, whether agents can prove to the entrant whether or not they contributed $e = 1$ in period 1. We model this as a Stackelberg problem. First, an incumbent commits to a matching mechanism for both periods. Then, after observing the incumbent's mechanism, an entrant commits to a matching mechanism for period 2 only. While platforms have commitment power agents do not and, in date 2, they can switch to the entrant's platform. We assume that there is no multi-homing and therefore that, in period 2, agents must either the incumbent's platform, the entrant's platform or their outside option.

In this game, there exist multiple equilibria, which strongly depend on agents' beliefs. In particular, for any strategies chosen by the entrant and the incumbent, there always exist an equilibrium where, in period-2, no agent joins the entrant's platform and one where no agent stays with the incumbent's platform. Therefore, we will focus on entry-proof mechanisms, defined below in Definition 4.

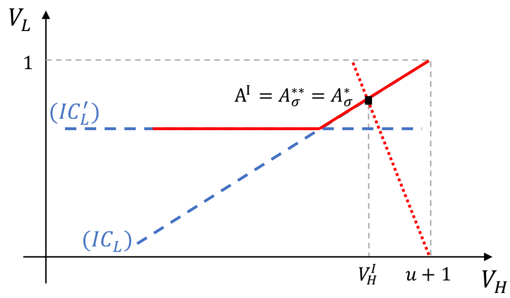
Definition 4. *(Entry-proofness). A mechanism implemented by the incumbent is said to be entry-proof if, of all possible equilibria, the one where all high-type agents remain on its platform delivers them the highest period-2 expected payoffs.*

Lemma 4. *The monopolist matching mechanism in the two-period game, ϕ_σ^{**}, is entry-proof in the regime where adverse selection dominates ($\sigma \geq \bar{\sigma}$) but not in the regime where moral hazard dominates ($\sigma \in [\bar{\sigma}, \overline{\overline{\sigma}}]$).*

Proof. When $\sigma \geq \overline{\overline{\sigma}}$ then, in period 2, ϕ_σ^{**} implements the second best static mechanism; that is the mechanism that maximizes the high-type agents' payoffs. Formally, (IC_L) binds. See Figure 9a. Therefore, by definition, the entrant cannot implement a mechanism providing strictly higher payoff to the high-type agents. As a result, the entrant cannot attract any agent on its platform.

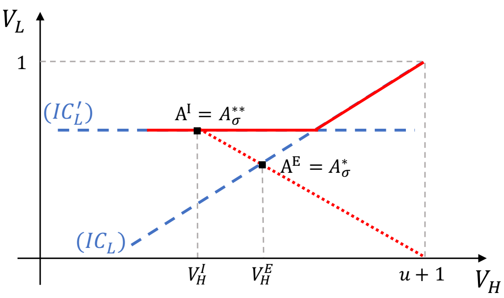
When $\sigma \in [\bar{\sigma}, \overline{\overline{\sigma}}]$ then we have seen that, compared to the static mechanism, ϕ_σ^{**} leaves rent to the low-type agents in period 2 to incentivize effort in period 1. Formally, when (IC_L') is satisfied, (IC_L) is slack. As a result, the entrant can offer the static mechanism which strictly increases the high-type agents' payoffs. See Figure 9b.

□



(a) Adverse selection dominates: $\sigma \geq \overline{\overline{\sigma}}$

Note. The incumbent's and the entrant's mechanism coincide with the optimal static mechanism ϕ^*_σ



(b) Moral hazard dominates: $\sigma \in [\overline{\sigma}, \overline{\overline{\sigma}}]$

Note. The incumbent's mechanism provides V_H^I to high-type agents in date-2 and the entrant's mechanism provides them $V_H^I > V_H^E$.

Figure 9 – Entry-proofness of the monopolist's matching mechanism in the two-period game

2.4.1 Entry-proofness without data portability

To make things interesting, assume that $\sigma \in [\overline{\sigma}, \overline{\overline{\sigma}}]$ so that the monopolist matching mechanism ϕ^{**}_σ is not entry-proof.

The entrant. The entrant takes the incumbent's matching mechanism as given and designs a matching mechanism in order to maximize welfare on its platform. There is no data portability therefore the entrant has no information about agents' date-1 reports about their types nor about whether they have contributed $e = 1$ or not. Therefore, designing the monopolist's static matching mechanism, ϕ^*_σ, is a dominant strategy for the entrant. In that case, if all agents switch to the entrant's platform in date 2, then high-type agents get payoff $V_H^E = V_H^*(\sigma)$.

The incumbent. The incumbent designs a matching mechanism to maximize the number of agents contributing $e = 1$ in period 1 and minimize the number of low-type agents on the platform in date 2. On top of the constraints from the monopoly situation, the incumbent now competes with the entrant in date 2. It is necessary that high-type agents get a greater payoff on the incumbent's platform than on the entrant's platform. This introduces the following new participation constraint

$$V_H^I \geq V_H^E \qquad (IR_H^I)$$

Under the entrant's matching mechanism, high-type agents get $V_H^*(\sigma)$ in period 2, which is independent of the incumbent's mechanism. Thus, the incumbent only needs to design the welfare maximizing mechanism providing $V_H^*(\sigma)$ to high-type agents in date 2.

Because under ϕ_σ^{**} we have that $V_H^{**}(\sigma) < V_H^*(\sigma)$ and $V_H^{**}(\sigma)$ increases with σ, it necessary to reject low-type agents to increase the share of high types in the population. However, under ϕ_σ^{**}, moral hazard dominates (ie.(IC_L') binds) and therefore it is necessary to allow some low-type agents to shirk in date 1 to exclude more in date 2.

Recall that welfare only depends on the number of high-type agents on the platform in each period and on the contributions of low-type agents in the first period. Therefore, because the incumbent accepts all high-type agents on its platform, its objective is to minimize the number of low-type agents allowed to shirk in date 1 to provide $V_H^*(\sigma)$ to high-type agents in date 2.

Optimal entry-proof mechanism. We now characterize the optimal entry-proof mechanism for the incumbent, without data portability.

First, it must be that (IC_L') binds; that is, low-type agents are indifferent between contributing $e = 1$ and having a chance to be on the platform in period 2 or not contributing and then being excluded. Otherwise, the incumbent can incentivize more low-type agents to contribute $e = 1$ in date 1.

Second, date-1 matches are only redistributive and therefore are irrelevant from the perspective of welfare. Without loss of generality, we can assume that the incumbent applies random matching in date 1. Therefore, the constraints of the incumbent's program are all independent of date-1 mechanism and therefore the analysis is the same as in the monopoly case with two periods. Thus, in the optimal entry-proof mechanism, low-type agents who contribute in date 1 are then matched according to $\phi_{\hat{\sigma}}^{**}$ in date 2, for some $\hat{\sigma}$ to be determined, and those who do not contribute are automatically excluded in period 2.

Fix σ, the share of high-type agents in the population. Define $\hat{\sigma} = \min_{\{} \sigma' | V_H^{**}(\sigma') \geq V_H^*(\sigma)\}$ and $\hat{\alpha} = \frac{(1-\hat{\sigma})\sigma}{(1-\sigma)\hat{\sigma}}$ the share of low-type agents that it is necessary to reject to obtain a population with a share $\hat{\sigma}$ of high-type agents when it was initially σ. See Figure 10. The optimal entry-proof mechanism, denoted $\hat{\phi}_\sigma$, is as follows. In period 1, the incumbent asks all agents to reveal their type, accepts all agents on the platform and matches them randomly.

A share $\hat{\alpha}$ of those who report being low type learn that they will be excluded in date 2, and therefore they choose effort $e = 0$ in date 1. All the other agents contribute $e = 1$ in period 1 and, in period 2, are matched according to $\phi^{**}_{\hat{\sigma}}$.

We summarize this characterization in Proposition 11 and provide the main consequence of interest for this paper in Corollary 1.

Proposition 11. *(entry-proof mechanism without data portability) The optimal entry-proof matching mechanism without data portability is $\hat{\phi}_{\sigma}$. To prevent entry, the incumbent allows some low-type agents to shirk in period 1 and excludes them in date 2. The incumbent applies the optimal matching mechanism with no entry to the rest of the agents.*

Corollary 1. *Threat of entry without data portability reduces effort provision in period 1 and increases exclusion in period 2.*

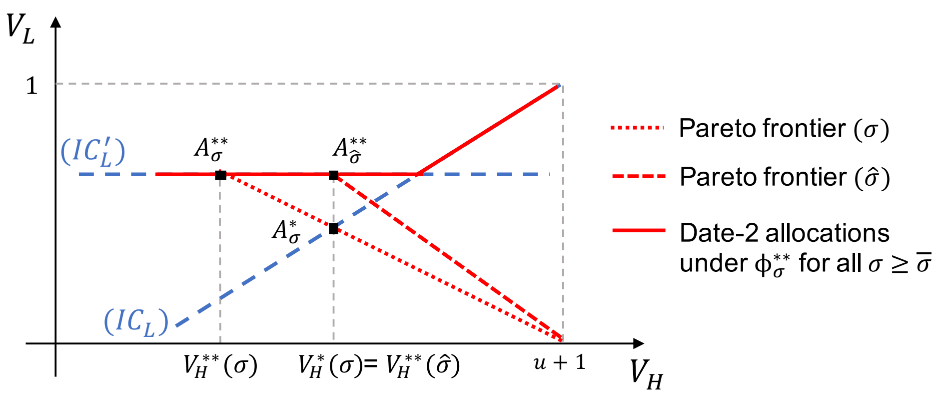


Figure 10 – Equilibrium in the case of period-2 entry without data portability.

Note. $A^{**}(\sigma)$ is the date-2 allocation under the monopolist matching mechanism in the two-period game, ϕ^{**}_{σ}. $A^{**}(\sigma)$ is the date-2 allocation under the entrant's matching mechanism, which coincides with the static monopolist mechanism, ϕ^{*}_{σ}. $A^{**}(\hat{\sigma})$ is the date-2 allocation under $\phi^{*}_{\hat{\sigma}}$, which is the incumbent's matching mechanism for the agents who have contributed $e = 1$ in the first period.

2.4.2 Entry-proofness with data portability

We now introduce data portability which we model by assuming that agent's can disclose their ratings to the entrant. Recall that, by assumption, agents truthfully rate their partner at the end of date 1. Therefore, all agents who have contributed $e = 1$ in period 1 obtain a good rating while the others obtain a bad rating. Because only low-type agents may receive a bad rating, those with a high rating have an incentive to disclose it to the entrant. Therefore, data portability and publicly observable actions are equivalent. As a result, with data portability, the entrant can identify the agents who have not contributed $e = 1$ and thus does not need to leave them rent to reveal their types.

To illustrate this point, suppose that the incumbent has designed the optimal entry-proof in the case with data portability $\hat{\phi}_\sigma$. In date 2, an entrant without data portability faces a population of size 1 with a share σ of high types. However, an entrant with data portability automatically rejects the agents who have not contributed in date 1 and thus faces a population of size $1-(1-\sigma)\hat{\alpha}$ with a share $\hat{\sigma} > \sigma$ of high types. Therefore, without data portability the highest payoff the entrant can provide to high-type agent is $V_H^*(\sigma)$ while it is $V_H^*(\hat{\sigma}) > V_H^*(\sigma)$ with data portability. See Figure 11.

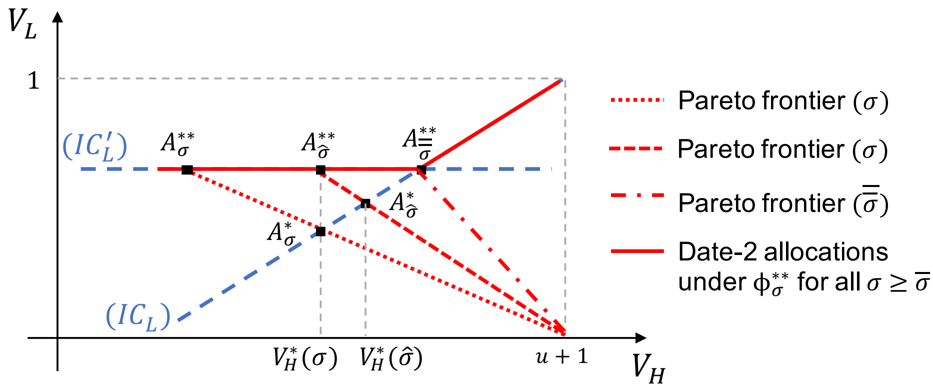


Figure 11 – Date-2 allocations in the case of period-2 entry with data portability.

Note. $A_{\overline{\overline{\sigma}}}^{**}$ is the allocation providing the highest payoff to high-type agents under an entry-proof mechanism when there is data portability. Under the optimal dynamic mechanism with no entry the date-2 allocation is A_σ^{**} which is not entry-proof because an entrant can implement A_σ^*. Under the optimal dynamic mechanism with entry but no data portability the date-2 allocation is $A_{\hat{\sigma}}^{**}$ which is not entry-proof because an entrant with data portability can implement $A_{\hat{\sigma}}^*$.

More generally, any matching mechanism such that (IC_L) is slack is not entry-proof, Lemma 5. This can be shown with a straightforward proof by contradiction, which we omit.

Lemma 5. *Consider the matching mechanism of an incumbent and assume that in the second period a second platform can enter and there is data portability. Under the incumbent's matching mechanism, if (IC_L) is slack then the matching mechanism is not entry-proof.*

The consequence of Lemma 5 is that the incumbent must exclude sufficiently many low-type agents at the end of date 1 so that such that (IC_L) binds. Building from the analysis without data portability, we know that the optimal way to do so is to allow a some low-type agents to shirk in period 1 and excludes them in period 2 and to apply the optimal matching mechanism with no entry to the rest of the agents.

Recall that $\overline{\overline{\sigma}}$ is defined such that the Pareto frontier passes through the point where (IC_L)

and (IC'_L) intersect; see Figure 11. Define $\overline{\overline{\alpha}} = \frac{(1-\overline{\overline{\sigma}})\sigma}{(1-\sigma)\overline{\overline{\sigma}}}$ the share of low-type agents that it is necessary to reject to obtain a population with a share $\overline{\overline{\sigma}}$ of high-type agents when it was initially σ. By construction, $\overline{\overline{\alpha}}$ is the minimum share of low-type agents that the incumbent must exclude to prevent entry when there is data portability.

Therefore, the following mechanism denoted $\overline{\overline{\phi}}_\sigma$ is the optimal entry-proof mechanism when there is data portability. In period 1, the incumbent asks all agents to reveal their type, accepts all agents on the platform and matches them randomly. A share $\overline{\overline{\alpha}}$ of those who report being low type learn that they will be excluded in date 2, and therefore they choose effort $e = 0$ in date 1. All the other agents contribute $e = 1$ in period 1 and, in period 2, are matched according to $\phi^{**}_{\overline{\overline{\sigma}}}$.

We conclude in Proposition 12, by comparing qualitatively the optimal entry-proof matching mechanisms with and without data portability.

Proposition 12. *When there is a threat of entry in the second period, data portability reduces effort provision in period 1 and increases exclusion in period 2.*

2.5 Conclusion

In this paper we have explored the idea that a platform can design a matching mechanism to solve issues stemming from adverse selection and moral hazard. The main take-aways are the following.

First, in the absence of transfers, positive assortative matching (PAM) is optimal in a static game with symmetric information only if there is types are public knowledge. Otherwise, the principal screen types by offering high quality - low quantity contracts to low-type agents and low quality - high quantity contracts to high-type agents. Furthermore, in a two-period game, further deviations from PAM may be necessary to incentivize low-type agents to exert effort in date 1.

Second, (threat of) entry and data portability both reduce date-1 effort provision and increase date-2 exclusion of low-type agents. An entrant does not internalize the effect of entry in date 2 on date 1 incentives and therefore entry creates a hold-up problem. Low-type agents who have exerted effort in date 1 are no longer compensated in date 2 if all migrate to the entrant's platform. To block entry, the incumbent must commit to screen more in date 2 even if it reduces effort provision in date 1. Data portability exacerbates the issue because it allows the entrant to screen talent even better.

The analysis of this paper opens several alleys for future research. First, some theoretical predictions may be tested empirically. In platforms of the gig economy, high-type agents would typically correspond to flexible users: such as drivers who can accommodate their passengers' demand (ride sharing) or hosts able to dedicate time to welcome their guests (lodging). The contracts setup by some platforms also share similar features with the screening contracts elicited in the paper. On Airbnb, hosts can choose between instant booking and manual booking, but hosts on instant booking appear higher up in the search list and thus have a higher probability to be selected by guests. On Uber, drivers can cancel rides but incur financial penalties and risk being excluded if they cancel too many. Thus there is a form of self selection of driver deciding to stay active on the app and those who disconnect more often. Regarding empirical estimations, it could be interesting to evaluate the effect of the entry of platforms in new cities on the incumbents' ratings policies. Also, in Europe, the GDPR now allows users to port personal data from any platform and our model predicts that it could affect the way ratings are computed.

Another promising application of the model is urban planning - although, the model should then allow agents to be horizontally differentiated as well. Governments sometimes intervene in urban planning to promote social and cultural diversity. The goal is to encourage different populations to interact which would enhance community harmony. Our model - and its variants with horizontal types - could be relevant because social interactions typically do not involve transfers and, instead, inclusion and ostracism are common practices. Theory predicts that a decentralized market would lead to PAM and that segregated communities would form - for instance, China towns, gay districts, gated communities, ghettos, etc. But our model shows that, in the absence of transferable utility, PAM may not provide sufficient incentivizes for pro-social behavior and social cohesion. Future research could test whether existing laws can be rationalized as attempts to use matching as a disciplining device and measure their impact empirically. One could study, for instance, the case of France where the state imposes by law 25% social housing in every town to avoid social segmentation and promote social diversity. Relatedly, in Denmark, the government has recently proposed a maximum of 30% 'non-western' residents in any neighborhood so at to avoid the "the risk of the emergence of parallel societies from a social and religious point of view".

Chapter 3

Barriers to real-time electricity pricing: Evidence from New Zealand

Abstract

We study the introduction of real-time electricity pricing in New Zealand and shed light on why adoption was low. Under this tariff, consumers are exposed to half-hourly varying spot prices, which are uncertain and volatile. We find that when spot prices spike, prospective adopters forego adoption. During a period of sustained high spot prices, recent adopters were substantially more likely to leave real-time pricing than more experienced consumers and were less likely to return once prices had fallen again. Our results suggest that that the timing of adoption matters and that providing information to consumers may help foster adoption, but that spot market crises leading to high electricity prices may pose a large threat to wide-spread adoption of real-time pricing.

This chapter is based on a project co-authored with Kevin Remmy

3.1 Introduction

Economic theory predicts that introducing real-time electricity pricing (RTP) in an economy with rational and perfectly informed agents will lead the retail market to unravel gradually. Under this tariff, consumers face spot prices and therefore pay the cost of their consumption in real-time rather than a price based on everyone's consumption and averaged over the year. Thus, the consumers with the consumption profiles least costly to serve self-select into RTP first, increasing the average cost of serving the other consumers. As a result, retailers increase their rates, and a new set of consumers finds it profitable to switch to RTP. This spiral between higher rates and more switching goes on until a significant share of consumers are on an RTP plan.[1] This scenario did not occur in New Zealand: since RTP was first introduced in 2013, the share of residential consumers on this tariff has remained below 1.25%. The purpose of this paper is to explain this puzzle and identify the barriers to wide-spread adoption of real-time pricing.

Because traditional tariffs do not reflect the varying cost of electricity, the demand of residential consumers has historically been inelastic. Increasing demand response could reduce the need to install generation capacity that is only used a few hours each year when demand peaks (Borenstein, 2005a), help integrate intermittent renewable energy sources such as wind and solar (Fabra, 2021), and reduce the incentives of producers to abuse their market power (Poletti and Wright, 2020). Thus, the prospect of significant efficiency gains stemming from increased demand response has led many economists and policy-makers to advocate for RTP (eg. Borenstein (2005a); Wolak (2013); Joskow and Wolfram (2012)). On the other hand, spot prices are uncertain and volatile and therefore RTP exposes consumers to the risk of a crisis on the spot market.[2]

Whether practitioners and policy-makers should promote RTP and foster its adoption is a hotly debated topic and different policies have already been implemented. In Europe, the Electricity Directive[3] ensures that all residential consumers have access to RTP contract by requiring large retailers to offer it by 2025. In Spain, the government has decided to default all residential consumers to RTP in 2015, while leaving them the possibility to opt-out and sign another contract with a retailer of their choice (eg. Fabra et al. (2021)). These situations

[1] See for instance Borenstein (2005b).

[2] For instance, the extreme winter storms that occurred in Texas in February 2021 led spot prices to spike and reach their regulatory ceiling of $9,000, jeopardizing the budgets of consumers who had signed contracts indexed to wholesale market prices. For an overview and a discussion of the events, see `https://www.tse-fr.eu/winter-texas`

[3] Directive (EU) 2019/944, June 2019

are respectively referred to as opt-in and opt-out policies. Our analysis of the introduction of RTP in New Zealand- which was a private initiative- is relevant both for opt-in and opt-out set-ups because it sheds light both on what drives adoption but also attrition.

Because the main specificity of real-time pricing is its dependence on uncertain and volatile spot prices, we are mostly interested in investigating how consumers react to spot pric spikes. In particular, we address the following questions. How do consumers react to spot price spikes? And how do these spikes affect attrition and adoption of real-time pricing contracts? To answer these question, we use a unique data set composed of all electricity retailer switches by residential households in New Zealand from January 2013 to June 2018. We also observe each household's monthly electricity consumption, half-hourly spot prices, and detailed census data that we can match to each household.

The fact that the retail market did not unravel generates a unique setting to study the adoption of real-time pricing. First, it allows us to study consumer choices between real-time pricing and more traditional tariffs over a long period with important price variations, including a sustained period of high spot prices. Second, we observe the first 4 years following the introduction of real-time pricing. As a consequence, there exist large differences in the time spent on RTP between the initial and the last adopters in our dataset. This feature allows us to study both how varying characteristics of adopters and time spent on real-time pricing affects attrition.

Following the adoption of RTP, nearly no consumers left RTP until a crisis on the electricity spot market during the winter of 2017 (hereafter referred to as the crisis). We find that during this crisis, the share of consumers switching to another tariff decreased with the time spent on RTP before the crisis. Furthermore, among consumers who discarded RTP during the crisis, the share of those switching back after the crisis increases with the time spent on RTP before the crisis. Exploiting temporal variations in the roll-out of real-time pricing across different cities, we rule out that this correlation is due to selection effects. The setting in New Zealand also allows us to study consumer choices between real-time pricing and other tariffs over a long period with important spot price variations. We build several natural predictors for future expected spot prices that consumers may consider and show that recent spot prices better explain adoption decisions. This result holds for predictors of average spot prices over the long run (one year), the medium run (three months), and the short-run (one month). We also find tevidence that most consumers choose to forego adoption rather than postponing it.

Overall, our findings suggest that inexperienced consumers - prospective and recent adopters - strongly react to spot price spikes, but that time spent on real-time pricing reduces reactivity. These findings may be a sign that consumers are present biased. Moreover, our results suggest that recent adopters of real-time pricing that experience a spot market crisis get scared and never return to real-time pricing. The combination of high spot price volatility starting with the Winter 2017 crisis coupled with present biased consumers could have led to a jamming of the unraveling process. Indeed, adoption rates are high before the crisis when spot prices are stable, and low following the crisis when there are large variations in spot prices. When spot prices spike unexpectedly adoption decreases and attrition increases - particularly for recent adopters.

First, our findings suggest that retailers or policy makers willing to foster the adoption of real-time pricing need to be "lucky" and hope that no unexpected period of high prices spikes arises. However, the fact that experience matters - in the sense that consumers become less sensitive to ongoing events after they have become familiar with the tariff - implies that timing matters. Strategically timing when consumers adopt (in an opt-in set-up) or are defaulted to (in an opt-out set-up) real-time pricing can increase the chances that consumers remain on real-time pricing.[4] However, the time horizon necessary for widespread adoption may be long: in New Zealand, adoption was suddenly interrupted by the winter 2017 crisis which happened more than three years after RTP was first introduced. Second, because present bias suggests that inexperienced consumers use inadequate information, either because computing long-term payoffs is costly or because they do not understand the spot price formation process. Either way, this observation calls for providing information to help consumers make rational and informed decisions. It is essential that they understand how spot prices form and that long-run gains can compensate immediate losses. Also, easing the access to consumption profiles recorded by smart meters can help consumers estimate whether it is beneficial for them to adopt real-time pricing. Finally, our results also suggest that spot market crises that lead to spiking electricity prices for consumers may pose a serious threat to wide-spread adoption of real-time pricing, as it may jam the unraveling process and scar consumers. To our knowledge, we provide the first evidence of how real-time pricing consumers behave during such a crisis. Due to a larger share of intermittent electricity generation sources and more prevalent extreme weather events due to climate change, spot market crises may become more

[4]In an opt-out set-up, the date when consumers are defaulted to RTP is a choice variable. In an opt-in set-up this is no longer the case because consumers decide whether or not to switch but one can time when to encourage switching, for instance through advertising campaigns or with subsidies.

common in the future.

Literature review. Our paper relates to two strands of the literature. First, it relates to the literature on dynamic pricing of electricity. On the theory side, Joskow and Tirole (2006) show that, in an economy with rational consumers, and in the absence of agency costs, real-time pricing implements the Ramsey optimum. Their result is satisfied even if consumers are imperfectly reactive to spot prices because paying attention is costly and they (rationally) choose their degrees of awareness. A direct implication of this result is that, in the absence of frictions, the retail market unravels until all consumers have adopted RTP in the steady state, with no need for intervention.[5] The low take-up of RTP in New Zealand contradicts this prediction and justifies our approach to identify frictions and departures from the rational consumer theory. The closest empirical papers to ours are Fowlie et al. (2020) and Ito et al. (2021) because they study consumer choices of dynamic tariffs. Both run randomized experiments and study adoption as well as consumption. Ito et al. (2021) document selection on price-elasticity and consumption patterns and shows that providing consumers with information about expected financial payoffs from switching can significantly increase adoption rates. Fowlie et al. (2020) compare the adoption rates and aggregate demand response under opt-in and opt-out set-ups. They find that demand response decreases over time among always-takers and increases over time among complacents. However, they consider dynamic tariffs where rates are set ex ante and therefore cannot address issues related to spot price uncertainty, which is a key element in the case of real-time pricing. In their set-ups, the only uncertainty consumers face relates to their preferences and in particular how costly it is to change their consumption habits. While they identify consumer learning, they find low rates of attrition. On the contrary, in the case of real-time pricing in New Zealand, we show that unexpected price spikes can lead to important attrition rates.

Second , our paper relates to the literature studying behavior that departs from the benchmark of rational and fully informed agents. Consumers may be present biased and rely on simple heuristics to make decisions with long term consequences. For instance, the weather can affect investments in solar panels (Lamp, 2018) or car purchases (Busse et al., 2015). Relatedly, Anderson et al. (2013) show that individuals often make "no-change" forecasts about gasoline prices. In the case of adoption real-time pricing of electricity, we show that recent or current

[5]The mechanism of unraveling, elicited for instance in Borenstein (2005b), is the following. The consumers with the consumption profiles least costly to serve self-select into RTP first, increasing the average cost of serving the other consumers. As a result, retailers increase their rates, and a new set of consumers finds it profitable to switch to RTP. This spiral between higher rates and more switching goes on until a significant share of consumers are on an RTP plan.

spot prices significantly affect consumer decisions. Furthermore, a growing literature shows that personal experience affects individuals' decisions. In macro-economics, Malmendier and Shen (2019) shows that experiencing periods of unemployment has long-term effects on consumption decisions. In finance, Hirshleifer et al. (2020) finds that analysts are biased by their first impressions. In industrial organization, Miravete (2003) show that consumers learn about their preferences after they have chosen a phone plan and make new choices accordingly. In the case of real-time electricity pricing, we show that consumers with bad first impressions may discard the tariff but that, with experience, consumers focus less on immediate outcomes.

The rest of the paper proceeds as follows. Section 3.2 describes our data and the context of our analysis. In Section 3.3 we study the behavior of consume on real-time pricing and, in particular, attrition and demand response. In Section 3.4 we study the decision-making process of prospective adopters of real-time electricity pricing. Section 3.6 discusses policy implications and Section 3.7 concludes.

3.2 Context and Data

3.2.1 The retail electricity market in New Zealand

New Zealand initiated the liberalization of electricity markets in the late 1980s, establishing competition in generation and retailing, while transport and distribution became regulated monopolies. Entry of new retailers remained limited for some time but eventually grew. While there were 15 retail companies at the end of 2013, 32 retailers at the beginning of 2018 offered contracts under 48 brands. Yet, the retail market remains dominated by historical incumbents, known as the 'Big 5', with a collective market share of around 90% between 2013 and 2018. Electricity is traded on a wholesale electricity market since the end of the 1990s. Some retailers - such as the 'Big 5' - are vertically integrated while others, generally the entrants, purchase electricity directly on the wholesale market.

Public initiatives encouraging consumers to switch have partly facilitated entry. An online price comparison tool, Powerswitch, was created by a consumer advocacy group with the government's support. Furthermore, in 2011 the regulator (Electricity Authority) started a campaign called "What's My Number" to inform and educate consumers about retail market opportunities. Daglish (2016) reports that this campaign increased switching rates significantly. The switching rate was about 20% on average annually between 2013 and 2018, which is greater than that of any European country.[6]

New Zealand has been a world leader in the deployment of smart electricity meters. By 2016 more than 50% of old meters were replaced by smart meters. The roll-out was based on private initiatives with no public intervention. While old meters measured the aggregate consumption and were read a few times a year, smart meters can measure and record electricity consumption in real-time. As a consequence, the roll-out of smart meters has allowed new electricity tariffs to emerge. Traditionally, consumers could only have flat tariffs - two-part tariffs with known fixed and variable components. Special electricity meters allowed time-of-use tariffs with, for instance, different day and night rates, but flat tariffs were and still are dominant. With smart meters, electricity tariffs can be more sophisticated. For instance, Electric Kiwi offers two-part tariffs and allows consumers to choose one hour of free consumption per day. Other retailers such as Paua to the People and Flick Electric offer real-time pricing tariffs under which consumers face the spot price of electricity which clears every half-hour.

3.2.2 Real-time pricing in New Zealand

In New Zealand, real-time pricing contracts were introduced by private retailers, with no public intervention[7]. To the best of our knowledge, only two retailers offered real-time pricing in the period covered in our dataset, Flick Electric Co. and Paua to the People. However, Flick Electric offered exclusively real-time pricing while Paua to the People also offered flat rates contracts.[8] Furthermore, most consumers adopting this tariff contracted with Flick Electric Co.[9] Therefore, in the rest of the paper, we focus exclusively on Flick Electric for real-time pricing tariffs and we know that a consumer joining Flick Electric adopts real-time pricing.

Flick Electric entered the retail electricity market for the first time at the end of 2013 in Wellington and then gradually entered other cities. Figure 12, below, shows that its market share initially grew quickly, stalled in June 2017 - the start of what we refer to as the winter 2017 crisis - and then remained slightly about 1%. At the end of May 2017, before the start of the crisis, Flick Electric's market share in New Zealand was 1.28% - or 23,057 households - with considerable heterogeneity across cities. While in Auckland, less than 1% of households chose real-time pricing, in Christchurch, they were 4.46%.

Tariff comparison. Often, electricity tariffs are two-part tariffs and consist of a fixed daily fee

[6] See Oxford Institute for Energy Studies (2019).

[7] In Spain, real-time pricing became the default tariff for all residential consumers in 2015. It is regulated, and consumers can opt-out if they prefer another tariff.

[8] See `https://www.rnz.co.nz/national/programmes/thiswayup/audio/201837850/power-to-the-people`

[9] In June 2017, Paua to the People had less than 1,000 customers, while Flick Electric had more than 23,000.

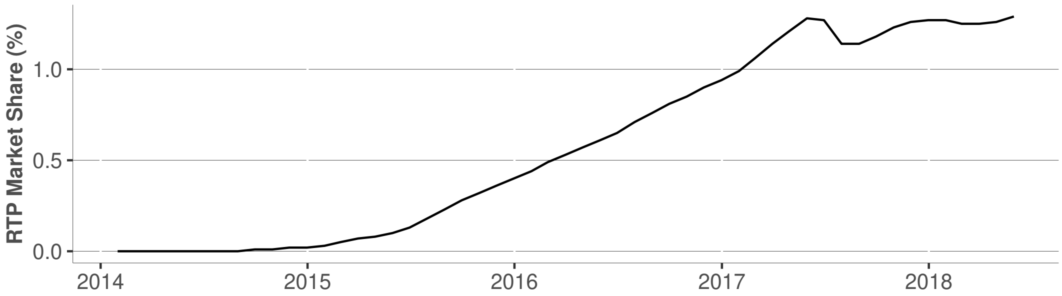


Figure 12 – History of the share of households on real-time electricity pricing and average monthly spot prices in New Zealand.

independent of consumption and a variable fee to pay per unit of electricity consumed. The real-time pricing tariff offered by Flick Electric is a two-part tariff, with the variable part being the sum of a pre-determined amount (to cover transportation and distribution plus a margin) and the spot price at the time of consumption. Thus, contrary to a flat tariff with a fixed variable part, the variable part under real-time pricing is uncertain and varies every half-hour. We illustrate this fact in Figure 13, where we plot the fixed and variable parts of real-time pricing offered by Flick Electric and of a flat tariff offered by Genesis Energy, the retailer with the largest market share in Wellington. Both Flick Electric and Genesis Energy adjusted their tariff only about once a year. However, spot prices vary substantially over time, so consumers on real-time pricing faced varying variable fees.

In Figure 13, both the flat tariff's fixed and variable parts always exceed that of real-time pricing, except during the winter 2017 crisis and at the beginning of the year 2018. We do not have data about consumer bills or which tariff they use when their retailer is no Flick Electric. Therefore we do not know how representative this comparison is. However, we have used a standard tariff offered by Genesis Energy, the retailer with the largest market share in Wellington, between 37% and 47% between 2013 and 2018. This difference is also coherent with the fact that consumers on a flat tariff pay a risk premium while those on real-time pricing face the risk directly. Furthermore, Flick Electric was a new entrant that needed to attract customers, while Genesis Energy was the incumbent with a large market share. Given the strong inertia in retail electricity markets, economic theory predicts that Flick Electric had an incentive to undercut its rivals. In contrast, Genesis Energy had little incentive to adjust its tariffs in response.[10] Finally, Flick Electric reports on its website[11] that their customers saved, on average, about 479NZ$ in 2017 compared to their previous tariff - while an average yearly

[10]See the extensive review about switching costs and consumer lock-in by Farrell and Klemperer (2007).
[11]Available at `https://news.flickelectric.co.nz/2017/08/01/the-inside-info-how-we-calculate-your-savings/`

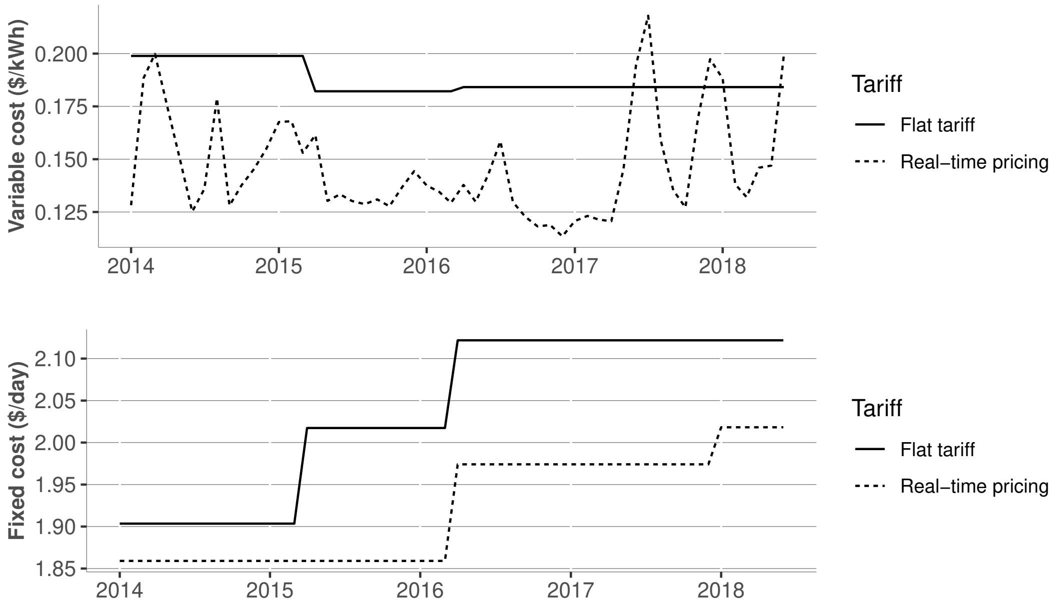


Figure 13 – Tariff comparison in Wellington.
Note: For the flat tariff, we use the tariff 'Household Composite (Standard user)' offered by Genesis Energy, the retailer with the largest market share in Wellington. For the real-time pricing plan, we use the 'Standard plan, All Inclusive' offered by Flick Electric. For the variable part of this tariff, we add the average monthly spot price of electricity to the variable part of the 'Standard plan, All Inclusive' of 2019 (for real-time pricing, apart from the spot prices, we do not have data about the variable part for the other years).

bill is about 2,200 NZ$/year.[12] For all these reasons, it seems plausible that during the period covered in our dataset (2013 to mid-2018), Flick Electric offered competitive tariffs compared with its rivals.

The winter 2017 crisis. The tariff comparison above highlights price variability and uncertainty as important features of real-time pricing. On top of these frequent price fluctuations, there is the risk that spot prices increase considerably. Such events happened several times in the period covered in our dataset. Our analysis will focus on one in particular that we refer to as the winter 2017 crisis. Studying a crisis on the wholesale market is interesting in itself because crises are inevitable. In consequence, practitioners and policy-makers need to understand how they affect consumers on real-time pricing. It is also an interesting event from an empirical strategy perspective because it affects everyone and thus reveals the state of mind of consumers at the same time. Specifically, during a crisis on the wholesale market,

[12] A similar magnitude was reported by Consumer NZ, in 2016: on average customers save 19% on their bills. See `https://www.energyawards.co.nz/finalist/2016/consumer-nz-energy-retailer-of-the-year/flick-electric-co-0`

the trade-off between short-term losses and potential long-term benefits is particularly salient to every consumer. The trade-off is salient to those who consider adopting real-time pricing and those who have already adopted it because there are no fees associated with switching to another tariff.

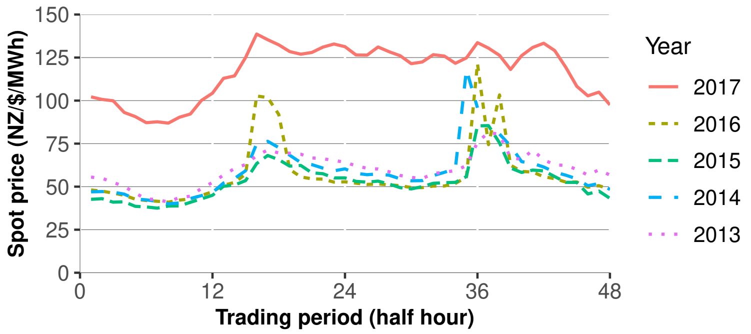


Figure 14 – Average half-hourly spot price in winter - from 2013 to 2017.

The winter 2017 crisis refers to a sustained period of high spot prices that occurred between June and August 2017. It is the first significant event on the electricity wholesale market since Flick Electric entered the market. About 16.2% of consumers on real-time pricing switched to another tariff during the crisis and 26.5% of them switched back to real-time pricing after the crisis. This crisis was due to low hydro levels coupled with high electricity demand.[13] Because about 60% of electricity comes from hydro generation in New Zealand, water shortage made electricity scarce, leading to high spot prices. As illustrated in Figure 14, spot prices increased two- to three-fold compared to previous winters. Such crises are rare - the last dry winter had occurred in 2008 - and its duration was hard to predict because it largely depends on rainfalls. While flat tariffs hedged consumers against the risk of a crisis, those on a real-time pricing tariff directly faced these high spot prices. Furthermore, spot prices during the winter 2017 crisis did not vary much throughout the day and only lowered during the night when consumers are asleep. Therefore, there was little room for consumers to adjust their consumption to avoid high prices. While we do not have information about consumer losses, Flick Electric reported that consumers made a loss of 80NZ$ from mid-June to mid-July compared with their previous tariff.

3.2.3 Data

We use a unique dataset containing all occurrences of consumers switching retailers between January 2013 and June 2018. These switches are recorded at the installation control point

[13] See Electricity Authority (2018)

(ICP)-level, a unique electricity meter identifier. We observe the previous retailer, the new retailer that the consumer is switching to, and the switching date. However, we do not observe which tariff the consumer chooses, only the retailer he contracts with. Furthermore, we observe whether the switch was related to the household moving into the accommodation or if it occurred while he was already living there.[14] Because we cannot trace where consumers move to nor which retailer a new tenant was with before moving in, we focus exclusively on switches unrelated to changing accommodation. Given that consumers often sign long-term contracts with their retailer that binds them even when they change accommodation, it is likely that our restriction is not too severe.

At the individual level (i.e. at the ICP-level), we also have yearly and monthly electricity consumption data. However, we do not observe consumers' half-hourly consumption used for billing them if they adopt real-time pricing. Furthermore, we observe the census tract in which an ICP is located, which allows us to merge the switching data to census data from 2013. We median data by census tract and use information on income, age, education, and work levels. A census tract usually contains between 50 and 80 households.

We collect publicly available aggregate data about each retailer's market share and the number of consumers each retailer gains and loses each month. Because this data is public, we can cover a longer period to the individual-level data we have.

We also collect publicly available spot price data for each network reporting region at the half-hourly level.[15] We use these half-hourly price data to compute the price faced by consumers on real-time pricing and compute average spot prices over different time horizons and locations.

Finally, we have information about the history of a subset of tariffs offered by each retailer in each network reporting region and their changes over time. In the dataset, the tariffs are two-part tariffs. For each tariff, we observe the fixed and variable parts, the prompt payment discount and electronic payment discounts, as well as the start and end dates at which these tariffs are available. However, we do not know whether a tariff is part of a long- or short-term contract or whether consumers have negotiated a discount.

In Table 1, we provide summary statistics of consumers. In addition, we compare the general population to consumers that switch retailers at least once during the sample period, and to consumers who switch to a real-time pricing contract during the sample period, in 2014 as well as in 2016. We can see that consumers who switch retailers have on average a

[14]We only observe those switches occurring due to moving where the retailer chosen by the new occupant is not the same as the retailer of the previous occupant.

[15]The electricity network in New Zealand is split into different network reporting regions (NRRs). The three largest cities, Auckland, Wellington, and Christchurch, belong to three distinct network reporting regions. We will focus our analysis on these three network reporting regions and refer to them by the name of the cities.

Table 1 – Summary statistics - Comparison of the median household, the median household switching retailer and the median household adopting real-time pricing - in Wellington.

	All ICPs		Switchers				RTP adopters			
			2014		2016		2014		2016	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Consumption (kWh/yr)	7.25	(3.8)	7	(3.8)	7.80	(4)	9.06	(4)	8.21	(3.8)
Income (NZ$/yr)	68.96	(29.6)	85	(31.6)	86.66	(31.7)	99.74	(29.2)	92.68	(30.7)
Age	39.55	(11.2)	36	(7.9)	36.76	(7.8)	37.15	(6.9)	36.47	(7.6)
Education (%)	18.62	(12.2)	28	(15.7)	28.98	(15.7)	36.66	(16.4)	32.15	(15.3)
Work (%)	54.32	(46.9)	50	(18.9)	48.90	(19.6)	55.32	(15.1)	52.20	(18.9)

higher income and a higher educational attainment compared to the general population. In comparison to both groups, consumers adopting real-time pricing have a higher electricity consumption, higher income, and a higher educational attainment on average. Note that the differences between consumers adopting real-time pricing compared to both switchers and the general population become smaller from 2014 to 2016.

3.3 Attrition of consumers on real-time pricing

In this section, we investigate the behavior of consumers who have adopted real-time pricing. In particular, we are interested in their response to variations in spot prices. We first investigate what drives the decisions of consumers who discard the tariff before investigating what drives consumer decisions to return to real-time pricing after having left.

3.3.1 The role of time spent on real-time pricing

To study these different effects, we focus on the winter 2017 crisis during which 19.7% of consumers on real-time pricing switched to another tariff. This event is relevant for our analysis for multiple reasons. It was a large and unexpected shock, and it affected all consumers on real-time pricing. Furthermore, while it was hard to anticipate, we argue that consumers were likely aware of the crisis once it occurred. Indeed, consumers are billed weekly, they receive notifications on their mobile app when prices spike, and their retailer, Flick Electric, regularly provided information about the crisis. Also, the event received media coverage. Therefore, all consumers on real-time pricing had to make a conscious decision to stay on or discard real-time pricing and, by a reveal preference argument, we can infer their preferences by studying their choices.

On Figure 15, we plot together the average spot price and then number of consumers on real-time pricing discarding the tariff each month. By and large, there is very little attrition unless spot prices spike, such as during the winter 2017 crisis. Between November 2013 and June 2017, only about 6.2% of all the consumers who had adopted real-time pricing eventually discarded it.[16] These low levels of attrition in the absence of spot price spikes suggest that consumers who discard real-time pricing do so in response to losses they are currently incurring.[17]

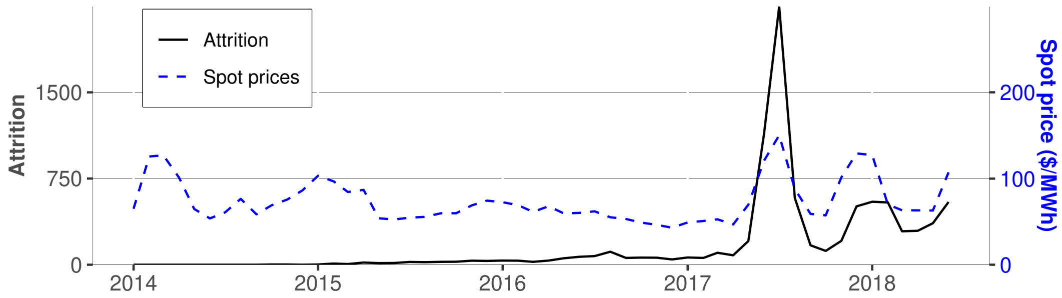


Figure 15 – History of monthly attrition from real-time pricing and average electricity spot prices.

A natural follow-on question is to understand why, when these price spikes occur, some consumers remain on real-time pricing while others switch to another tariff. While consumers with the largest electricity consumption incur more considerable losses when prices are high, they obtain larger benefits from real-time pricing when prices are low, relative to a flat tariff. Also, because real-time pricing is a new form of tariff and the formation of spot prices is a complex process, switching decisions may depend on how sophisticated consumers are - measured by socio-economic characteristics such as education - or how sophisticated they have become with experience. On Figure 16, we plot the share of consumers discarding real-time pricing as a function of the time they have spent on the tariff and compare consumers residing in the regions of Wellington, Christchurch and Auckland. The attrition rates are highest in Christchurch and lowest in Auckland, but the pattern is very similar in all three regions. The share of consumers discarding real-time pricing during the winter 2017 crisis decreases with the time spent on the tariff.

[16]To compute this number, we take the ratio between the number of consumers who discard real-time pricing for reasons unrelated to moving-in a new accommodation over the total net number of consumers who adopted real-time pricing.

[17]This is coherent with the literature on bill shock that finds that, in different retail markets, consumers switch tariffs after receiving exceptionally high bills. See, for instance, Grubb and Osborne (2015) for an analysis of bill shock in the cellular service.

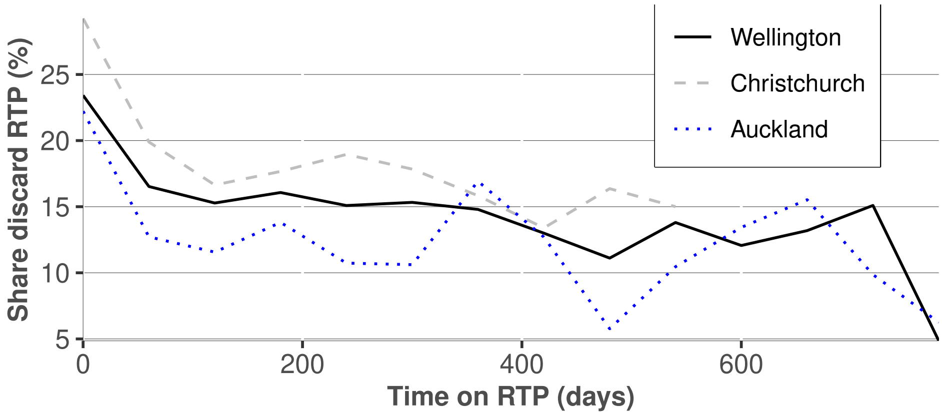


Figure 16

To investigate the effect that time spent on RTP had on leaving real-time pricing during the crisis, we regress the decision to discard real-time pricing during the crisis on the time spent on RTP, the winter electricity consumption and control variables. More specifically, we estimate the following equation:

$$\text{Discard RTP}_i = \alpha \text{Time on RTP}_i + \gamma \text{Winter Consumption}_i + X_i' \beta + \varepsilon_i, \tag{3.1}$$

where Discard $\text{RTP}_i \in \{0, 1\}$ is an indicator equal to 1 if consumer i decides to discard RTP during the winter 2017 crisis, Time on RTP_i is the number of months that consumer i spent on RTP prior to June 1st, 2017, Winter Consumption is the electricity consumption from June to August 2016, X_i contains control variables, and ε_i is assumed to follow a logistic distribution. In X_i, we control for census-level logged median household income, age, and work and education indeces, and consumer's previous retailer by location.

Table 2 – Discarding real-time pricing during the winter 2017 crisis.

	Dependent variable:					
	Discard RTP					
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP(month)	-0.038^{***}	-0.036^{***}	-0.021^{***}			
	(0.005)	(0.005)	(0.005)			
log(Time on RTP)				-0.229^{***}	-0.219^{***}	-0.140^{***}
				(0.023)	(0.024)	(0.038)
Joined Last			0.455^{***}			0.305^{***}
			(0.082)			(0.114)
Winter Consumption (MWh)		0.110^{**}	0.112^{**}		0.109^{**}	0.109^{**}
		(0.054)	(0.054)		(0.054)	(0.054)
log(Income) (k$/yr)		0.058	0.079		0.071	0.079
		(0.100)	(0.100)		(0.100)	(0.101)
Age		-0.010^{**}	-0.010^{**}		-0.010^{**}	-0.010^{**}
		(0.004)	(0.004)		(0.004)	(0.004)
Work (%)		−0.005	-0.005^{*}		-0.005^{*}	-0.006^{*}
		(0.003)	(0.003)		(0.003)	(0.003)
Education (%)		−0.003	−0.003		−0.003	−0.003
		(0.004)	(0.004)		(0.004)	(0.004)
Location-on-Previous retailer FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,703	8,394	8,394	8,703	8,394	8,394
Note:						*p<0.1; **p<0.05; ***p<0.01

The results are summarized in Table 2. The specifications in the first three columns assume a linear effect Time on RTP_i. In the last three columns, we use $log(\text{Time on RTP}_i)$. The effect of Winter Consumption is similar across all specifications, and statistically significant at the 5% level. Interestingly, the coefficient for time spent on the tariff is statistically and economically significant, both in the linear and the log specifications. Using the results from column 2, at average value of the covariates, spending 4 more months on real-time pricing decreases the probability to discard real-time pricing by 2.19 percentage points (the average probability to discard RTP is 19.75%).

We can also see that the coefficient on *Winter Consumption* is stiatically significant and positive. However, the economic effect is negligible. Using the results from column 2, at average value of the covariates, increasing Winter Consumption by 20% (or 210.49MWh), increases the probability to discard real-time pricing by 0.28 percentage points - which is low compared to 19.75%,the unconditional probability to discard it.[18]

The effects of consumer demographics are similar in all specifications and none are economically significant. Furthermore, $log(Income)$ and $Education$ are not statistically significant while Age is only significant at the 5% level and $Work$ at the 10% level. The fact that income is

[18]For robustness, we use two different specifications for electricity consumption - Annual Consumption (measured in 2015) and Seasonal Difference (measured as the difference between winter and summer consumption in 2016). The latter specification may be relevant if consumers react to bill shocks. In both cases we also conclude that the effect of consumption is not statistically significant. The results are in Table 9, in the Appendix.

small and not significant suggests that consumer decision to discard real-time pricing was not driven by wealth effects - which is consistent with the fact that households who adopt RTP generally have high income. Also, the fact education is not significant indicates that consumer sophistication did not play a role either.

In addition, we run the same regressions and control for a "first impression" effect for consumers who joined last with a dummy ('Joined Last') equal to one if the consumer adopted RTP with the last cohort. The goal is to ensure that the effect of time spent on RTP is not solely driven by the last adopters. The results are displayed in columns (3) and (6). In both cases, the coefficient for 'Joined Last', which means that the last consumers to join are significantly more likely to discard RTP. The effect of time spent on RTP diminishes by about one third but it remains statistically significant, both when measured in levels and in log terms.

We know that 26.55% of consumers who discarded RTP eventually switched back to it after prices resume to normal levels. However, we do not know if those consumers decided ex ante to arbitrage between tariffs or if they simply discarded RTP and later reconsidered their decision. Either way, the fact that they switch back to RTP means that the crisis did not traumatize them. In that respect, these consumers may be different than those who discard real-time pricing and do not return. To make sure that our results are not biased by these consumers, we run two types of robustness checks. We first run the same regressions as that the specification in column (3) and (6) of Table 2 except that we treat consumers who leave and come back as if they didn't leave at all. The underlying assumption is that these consumers had anticipated that they would switch back to RTP after the crisis. We also run the same regressions as that the specification in column (3) and (6) of Table 2 except that we remove consumers who leave and come back from the dataset and test what affects the decision to leave.

The results are summarized in Table 10, in the Appendix. In all cases, the effect of Time on RTP is greater than in the baseline regressions, which suggests that our results are robust.

3.3.2 Selection versus experience

In the next step we investigate what the coefficient *Time on RTP* captures. Two different mechanisms can explain why consumers who have spent more time on real-time pricing are more likely to remain on the tariff during the crisis.

The first possible mechanism is selection on unobservable characteristics at the time of adoption that play a role in the attrition process. Note that the results in Table 2 provide evidence that selection on observable characteristics did not play a role. Even when controlling for demographics and electricity consumption, time spent on RTP remains statistically and economically significant and of close magnitude - see columns (1)-(2) and (4)-(5) of Table 2. However, there may be unobservable characteristics uncorrelated with observable ones that explain the correlation between time spent on RTP and the decision to discard RTP during the crisis.

The second mechanism is an experience effect that makes consumers who have spent more time on real-time pricing less likely to leave during spot price surges. This effect could manifest itself through learning more about real-time pricing contracts and the formation of spot prices, for instance.

To shed light on which mechanism can explain the effect of time spent on real-time pricing, we take advantage of the fact that the tariff became available at different times in different cities. We reason by contradiction: Assume that some innate unobservable characteristics matter in consumer switching decisions during the crisis and that there was selection at adoption on this unobservable characteristic. Then, because the tariff was available in Christchurch about 22 months after its introduction in Wellington, we would expect that two consumers adopting real-time pricing at the same time in Wellington and Christchurch make two different switching decisions. If, on the other hand, we find no significant difference in the switching decision of two consumers adopting at the same time in Wellington and Christchurch, we can rule out that selection on unobservable characteristics plays a role.

Figure 17 shows that when real-time pricing became available in Christchurch, 39% of all consumers in Wellington who adopted RTP before the crisis had already adopted it. Then, under the (untestable) assumption that the selection of unobservable characteristics was similar in Christchurch and Wellington, we can conclude that the adopters in Wellington are significantly different from the consumers adopting real-time pricing in Christchurch at the same time. Therefore, we have a relevant set-up to test whether selection at adoption explains the correlation between time spent on real-time pricing and the decision to discard real-time pricing during the crisis.

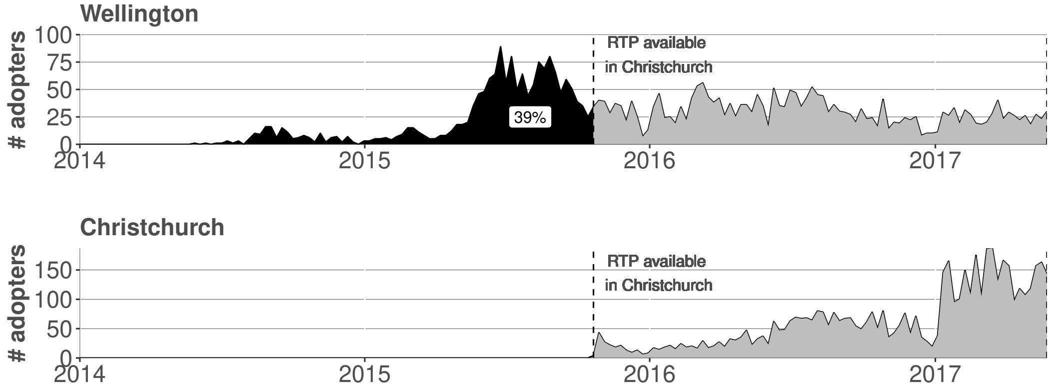


Figure 17 – Number of consumers adopting real-time electricity pricing every week in Wellington (top) and Christchurch (bottom) between November 2013 and June 2017.

Formally, we regress consumer decisions to discard RTP during the crisis (Discard $\text{RTP}_i \in \{0, 1\}$) on time spent on real-time pricing, a location dummy for Christchurch, an interaction between the location dummy and control variables:

$$\begin{aligned}\text{Discard RTP}_i =& \alpha_{exp}\text{Time on RTP}_i + \alpha_{loc}\text{Christchurch}_i + \gamma\text{Time on RTP}_i \times \text{Christchurch}_i \\ &+ X_i'\beta + \varepsilon_i,\end{aligned}$$

where Discard RTP_i and Time on RTP are defined as in 3.1, Christchurch is a dummy equal to one if the consumer in question lives in Christchurch, X holds consumer characteristics, and ε_i is a logistic error term. Our sample is the set of consumers who adopted real-time pricing in Wellington and Christchurch only after the tariff was available in Christchurch in September 2015. Therefore, the Wellington sample is truncated - we have removed the initial adopters - while the Christchurch sample is not. Our main variable of interest is γ. If selection at adoption were an important driver, the interaction variable γ between the location dummy and experience would be statistically significant.

The results are in Table 3. We see that across all specifications, γ is statistically - and economically - insignificant. This result suggests then that selection at adoption does not explain the correlation between time spent on the tariff and the decision to stay or opt-out during the winter 2017 crisis. For robustness, we repeat the same exercise between Auckland and Christchurch and find that the interaction term is insignificant as well. The results are in Table 11 in Appendix 3.C.3. The results of this section then lead us to conclude that consumers' perception of real-time electricity tariffs is affected by their experience with the tariff. As a result, the probability that a consumer on real-time pricing switches to another tariff during the winter 2017 crisis decreases with her experience.

Table 3 – Comparison of the probability to discard RTP in Wellington and Christchurch.

	Dependent variable:					
	Discard RTP					
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (month)	-0.045^{***}	-0.041^{***}	-0.046^{***}			
	(0.012)	(0.012)	(0.013)			
log(Time on RTP)				-0.224^{***}	-0.204^{***}	-0.221^{***}
				(0.058)	(0.062)	(0.068)
Christchurch	0.151	0.119	0.239	0.043	0.006	0.063
	(0.143)	(0.145)	(0.340)	(0.345)	(0.346)	(0.550)
Time on RTP x Christchurch	-0.0005	0.007	0.015			
	(0.015)	(0.015)	(0.017)			
log(Time on RTP) x Christchurch				0.026	0.036	0.069
				(0.066)	(0.066)	(0.084)
Yearly Consumption (MWh)	-0.019^{**}	-0.018^{**}	-0.018^{**}	-0.019^{**}	-0.018^{**}	-0.018^{**}
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Seasonal Difference (MWh)	0.384^{***}	0.377^{***}	0.379^{***}	0.383^{***}	0.378^{***}	0.381^{***}
	(0.080)	(0.081)	(0.081)	(0.081)	(0.081)	(0.081)
Income (k$/yr)	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Age	-0.009^{*}	-0.009^{**}	-0.009^{*}	-0.009^{**}	-0.009^{**}	-0.009^{**}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Work (%)	-0.007^{*}	-0.007^{*}	-0.007^{*}	-0.007^{*}	-0.007^{*}	-0.007^{*}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Education (%)	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Month FE?	No	Yes	No	No	Yes	No
Month-on-NRR FE?	No	No	Yes	No	No	Yes
Observations	6,104	6,104	6,104	6,104	6,104	6,104
Log Likelihood	$-3{,}086.962$	$-3{,}076.094$	$-3{,}072.237$	$-3{,}080.982$	$-3{,}076.097$	$-3{,}072.225$
Akaike Inf. Crit.	6,193.924	6,194.188	6,208.475	6,181.964	6,194.194	6,208.450
Note:						*p<0.1; **p<0.05; ***p<0.01

The role of inattention

Another reason why the time spent on RTP affects the decision of leaving RTP during the crisis could be inattention that sets in once consumers have spent a certain amount of time on the tariff. However, there are several reasons why we do not think inattention was an important factor: First, the price spikes during the winter of 2017 were very unusual compared to previous years and covered in media. Second, through a mobile phone app, Flick customers receive weekly updates on their bills and also receive alerts when prices are high. Both reasons make it unlikely that consumers were completely unaware of the crisis.

3.3.3 Switching back to RTP

Overall, 26.55% of consumers who discarded RTP during the winter 2017 crisis switched back to RTP after the crisis. In Figure 18 we plot the share of consumers who switched back to RTP as a function of the time they have spent with the tariff prior to the crisis. The probability to switch back to RTP increases from 15% for consumers who spent less than 100 days on RTP before the crisis to more than 30% for those who spent more than 500 days. We confirm this graphical evidence by regressing (logit) consumers decision de return to real-time pricing on the time spent on RTP and control variables, see Table 5 in Appendix 3.B. We find that time spent on RTP affects the decision to return to real-time pricing significantly, both statistically and economically. At the average of the covariates, increasing experience by 4 months increase the probability to return to RTP after the crisis by 4.8 percentage points. A revealed preference argument would suggest that only consumers with a good perception of the tariff would return, reinforcing the previous finding that consumer perception of real-time pricing improved with experience.

Reciprocally, this finding suggests that consumers who adopted shortly before the crisis started got scarred by a bad first impression and, accordingly, were more likely to discard the tariff forever. This result suggests that spot market crises leading to price spikes may permanently drive consumers away from choosing real-time pricing. Given that extreme weather events are likely to increase both in frequency and in magnitude, such price spikes may occur more often in the future, creating an important challenge for policymakers trying to achieve wide-spread adoption of real-time pricing.

The findings of this section also pose the question of whether inexperienced consumers over-reacted or whether experienced consumers under-reacted. Without having a clear benchmark of what the optimal behavior should have been, it is impossible to answer this question. We attempt to shed some light on whether inexperienced consumers tend to over-react to ongoing events in the next section.

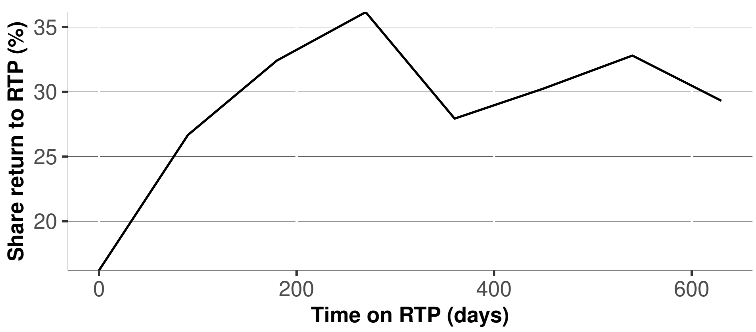


Figure 18 – Share of households switching back to RTP after opting-out during the winter 2017 crisis within three months after the end of the crisis as a function of the number of days spent on RTP prior to the crisis.

3.4 Adoption of real-time electricity pricing

Given that prospective adopters of real-time pricing are not experienced by definition, we attempt to determine whether they are sensitive to ongoing events. To do so, we study the role of contamporaneous spot prices on adoption of real-time pricing contracts. In the previous section, we have shown that experience matters in the sense that recent adopters react much more strongly during a crisis on the spot market than experienced consumers. In this section, we study whether inexperienced consumers are also sensitive to current spot prices and whether they over-react to short-run spot price fluctuations. In particular, our focus lies on the role of contamporaneous spot prices and whether consumers strategically time their adoption decisions. More generally, we are also interested in better understanding the decision-making process of prospective adopters.

3.4.1 Descriptive evidence

In Figure 19, we plot jointly the history of monthly spot prices and the number of consumers adopting real-time pricing in Wellington. Adoption rates correlate negatively with contemporaneous spot prices: when spot prices are high, adoption is low, and vice-versa. This relationship is particularly evident when the spot prices increase or decrease significantly, such as in 2017, when adoption is high both before and after the winter 2017 crisis but low during the crisis. The link between adoption and spot prices is also apparent for more minor variations, such as the period between late 2015 to early 2016.

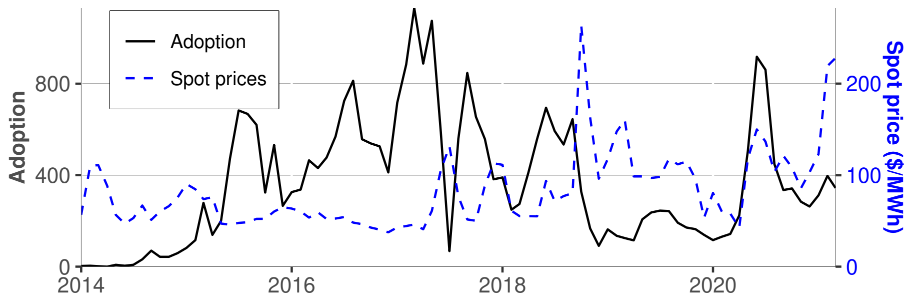


Figure 19 – History of monthly RTP adoption and average spot prices.

We provide additional graphical evidence on Figure 20 by plotting the share of households switching to a non-Big-5 retailer choosing real-time pricing as a function of the average spot prices in the four weeks preceding the switch. Again, the plot suggests that consumers are sensitive to prices contemporaneous to their switching decisions. The share of switchers adopting real-time pricing drops nearly 50% when spot prices are in the range 40-60 $/MWh to less than 20% when prices exceed 100 $/MWh.

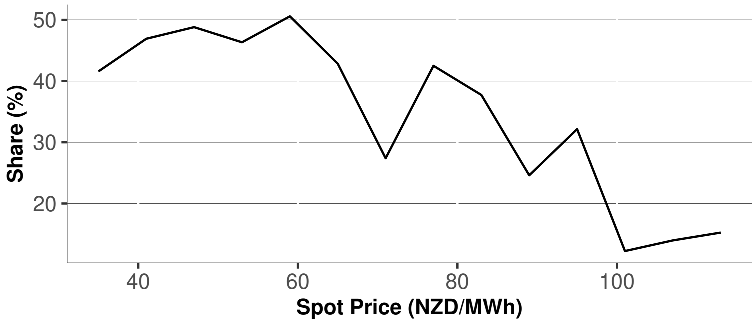


Figure 20 – Share of consumers switching to a non-Big-5 retailer who adopt RTP as a function of the average spot prices in the 4 weeks preceeding the switch - in Wellington.

The fact that prospective adopters react to spot prices contemporaneous to their switching decisions does not necessarily mean that they are not forward looking, especially if spot prices are persistent and can serve as a relevant proxy for future ones. In Table 6 in Appendix , we present the correlation between recent and future spot prices over several time horizons. We see that future prices do not correlate strongly with different definitions of recent prices. The definition for recent price that correlates most with future ones is the one computed over a two-week horizon. However, even under this definition, the correlation is small for time horizons greater than one month.

3.4.2 Empirical strategy

We now investigate more formally whether consumers react to recent spot prices rather than different proxies of future payoffs. In particular, we build several definitions of future prices that consumers may consider and investigate whether these future prices affect adoption decisions more than recent spot prices. Ideally, we would like to investigate this question by analyzing the decision-making process of all consumers. Due to data limitations, this is not possible. For instance, we do not observe whether consumers are locked into a tariff with high early-termination fees or not. In addition, some consumers may exhibit high inertia in their decisions (see Hortaçsu et al., 2017). Not accounting for this inertia would risk biasing our results. In order to circumvent these issues, we restrict ourselves to the subset of consumers who decided to switch. This restriction comes with some limitations itself. Real-time pricing attracts two types of consumers: those would have switched retailer anyway (business-stealing) and those who would only switch to real-time pricing (demand-creation). By restricting attention to consumers who switch retailers, we are not capturing the demand-creation effect, which may be important given that real-time pricing is a new, very different type of electricity tariff. However, real-time pricing was introduced in New Zealand at the same time as other retailers entered the market and offered, among others, various non-traditional tariffs.[19] Thus, in order to limit the bias of our estimates, we will restrict attention on the subset of consumers who switch to a retailer which is not an incumbent[20] assuming that, in that subset, the demand-creation effect is negligible.

We employ a logit model, where we regress the individual decision to adopt RTP upon switching retailers on the different price definitions, controls, and fixed effects. We build several definitions of future prices that consumers may consider and investigate whether, controlling for recent spot prices, future prices affect their adoption decisions. For each day, we define *Recent Price* as the average spot prices over the past four weeks ending that day, between 7am and 10pm.[21] We consider three proxies for*Future Price*: realized spot price, last year's price, and future price predicted based on an AR(1) process. Essentially, our three price definitions assume, respectively, perfect foresight, backward-looking behavior, and consumers acting as "econometricians" using a prediction model to make forecasts

Because spot prices are seasonal, we assume that the relevant benchmark for a rational forward-looking consumers is to look over a one-year period. However, forecasting spot prices over one year is a complicated exercise and forecasts are less reliable the further away in time they are. Therefore, it could be that consumers, while being forward-looking, do not trust their long-run forecasts and chose shorter a time horizon. Therefore, we compute for robustness future prices over 1-month ("short run") and 6-month ("medium run") horizons as well, next to the 12-month ("long run") horizon.

Formally, our specification writes

$$Y_{it} = \alpha_1 P_{mt,\text{Recent}} + \alpha_2 P_{mt,\text{Future}_f} + X_{it}\beta + \gamma_m + \lambda_t + \varepsilon_{it}, \tag{3.2}$$

where Y_{it} is equal to one if consumer i in market m at date t decides to adopt RTP (conditional on switching to a non-traditional retailer), $P_{mt,\text{Recent}}$ is the recent spot price, P_{mt,Future_f} is the future price, where $\text{Future}_f \in \{\text{realized, last year, AR(1)}\}$, X_{it} holds control variables and ε_{it} is assumed to follow a logistic distribution. In X_{it}, we control for yearly household consumption as well as consumption differences between winter and summer, the origin retailer, the winter 2017 crisis, the time after the 2017 winter crisis, and census-level median household income, age, and work- and education status.

[19]For instance, Electric Kiwi offered a flat two-part tariff and the customer could choose one hour per day when electricity is free.

[20]The Big 5 account for about 90% market share.

[21]We use this time horizon because consumers consume the bulk of their electricity during these hours.

Table 4 – Logit regression of switchers to non-traditional tariffs. Recent price always computed over 4 weeks. We use data from 2014-06-01 to 2018-06-01

	Dependent variable:								
	Individual decision to adopt RTP								
	1 Month	1 Month	1 Month	6 Months	6 Months	6 Months	12 Months	12 Months	12 Months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Recent Price	-0.171^{***}	-0.152^{***}	-0.149^{***}	-0.135^{***}	-0.151^{***}	-0.149^{***}	-0.132^{***}	-0.182^{***}	-0.149^{***}
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Future Price (realized)	0.040^{***}			0.091^{***}			0.381^{***}		
	(0.009)			(0.016)			(0.023)		
Future Price (last year)		0.013			-0.080^{***}			-0.135^{***}	
		(0.029)			(0.010)			(0.011)	
Future Price (AR (1))			0.021^{***}			0.066^{***}			0.131^{***}
			(0.007)			(0.024)			(0.047)
Consumption (MWh)	0.037^{***}	0.037^{***}	0.037^{***}	0.036^{***}	0.037^{***}	0.037^{***}	0.036^{***}	0.036^{***}	0.037^{***}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Seasonal Difference (MWh)	0.015	0.015	0.013	0.017	0.014	0.013	0.016	0.012	0.013
	(0.052)	(0.052)	(0.052)	(0.053)	(0.053)	(0.052)	(0.053)	(0.053)	(0.052)
Income	0.004^{***}	0.004^{***}	0.004^{***}	0.004^{***}	0.004^{***}	0.004^{***}	0.003^{***}	0.003^{***}	0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.006^{***}	0.006^{***}	0.006^{***}	0.007^{***}	0.006^{***}	0.006^{***}	0.006^{***}	0.006^{***}	0.006^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
White collar worker	0.621^{***}	0.619^{***}	0.618^{***}	0.623^{***}	0.618^{***}	0.619^{***}	0.573^{***}	0.609^{***}	0.619^{***}
	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)	(0.148)
Education	1.344^{***}	1.342^{***}	1.343^{***}	1.345^{***}	1.331^{***}	1.341^{***}	1.397^{***}	1.343^{***}	1.341^{***}
	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)	(0.162)	(0.161)	(0.161)
Winter Crisis	-1.407^{***}	-1.302^{***}	-1.365^{***}	-1.553^{***}	-0.871^{***}	-1.355^{***}	-1.100^{***}	-0.661^{***}	-1.355^{***}
	(0.111)	(0.108)	(0.110)	(0.117)	(0.121)	(0.110)	(0.109)	(0.121)	(0.110)
Post Crisis	-0.761^{***}	-0.691^{***}	-0.721^{***}	-0.799^{***}	-0.186^{*}	-0.718^{***}	-0.733^{***}	-0.242^{***}	-0.718^{***}
	(0.084)	(0.082)	(0.083)	(0.084)	(0.104)	(0.083)	(0.083)	(0.091)	(0.083)
Location FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34,000	34,000	34,000	34,000	34,000	34,000	34,000	34,000	34,000
Note:									*p<0.1; **p<0.05; ***p<0.01

The results are in Table 4. We can see that across definitions of future prices and across the time horizon over which we compute these future prices, *Recent Price* is always statistically and economically significant and keeps the same magnitude. In fact, the coefficient on *Recent Price* in column 1 of Table 4 means that at average value of the covariates, an increase in the spot price by one standard deviation decreases the probability of adopting real-time pricing by around 8.94 percentage points. For reference, at the average value of covariates, the probability of adoption is around 32.82%. Moreover, we see that the different definitions of future prices are either not statistically significant, have an unintuitive sign, or both. The only exceptions are *Future Price (last year)* over 6 and 12 months. These results suggest that consumers focus on contemporaneous prices rather than trying to predict long-run prices and hence the existence of present bias.

We also check whether the volatility of spot prices may play a role in adoption decisions. We compute the daily standard deviation of peak-hour spot prices and take averages over the last 1, 2, and 4 weeks. We then regress the individual switching decisions on both *Recent Price* and our measure of volatility, including all controls we use in (3.2). The results are in Table 7 in Appendix 3.B. We see that price volatility is not significant, leading us to conclude it does not play an important role in adoption decisions. As additional robustness checks, we re-run the specification (3.2) but make two changes: In the first robustness check, we change the time horizon over which we compute *Recent Price* to two weeks. In the second check, we keep the time horizon for computing *Recent Price* at 4 weeks, but include the log of prices (both recent and future) in the regressions. Our findings are robust to these changes. Detailed results for the robustness checks are in Appendix 3.D.

Do households forego or postpone adoption?

The results from the previous section suggest that households react to spot prices that are contemporaneous to their adoption decision. A natural follow-up question is whether households postpone or forgo adoption when spot prices are high. To answer this question, we focus on the winter 2017 crisis when spot prices surged and remained high for several weeks. Because spot prices more than doubled, consumers who were willing to adopt RTP and were able to postpone adoption had an interest in doing so.

In Figure 21 we plot the weekly number of households switching to real-time pricing for the first time in Wellington.[22] While the number of new adopters is relatively constant before the

[22] As we saw in the previous section, some consumers who were on real-time pricing switched to another tariff during the crisis and then switched back to real-time pricing. However, we do not consider them in our analysis. In the current analysis, we are only interested in the decision of consumers who were considering adopting real-time pricing for the first time.

crisis, adoption drops during the crisis with only about 25 new households per week joining. Interestingly, there is a surge in adoption in mid-August when spot prices start declining. After this surge, the number of households adopting real-time pricing for the first time becomes constant over time below the pre-crisis level, with about 16 new households per week. This surge may suggest that some consumers who would have adopted during the crisis waited until spot prices decreased. The goal then is to quantify whether waiting for the right price is a widespread strategy among prospective adopters or if only a few of them do. To do so, we need to predict how many consumers decided not to adopt during the crisis because of the crisis and compare it to the number of consumers who postponed adoption and adopted after prices decreased.[23]

To compute the number of consumers who decided not to adopt during the crisis because of the crisis, we predict the total number of consumers who would have adopted had there not been a crisis using a fixed-effects regression[24] and subtract the observed number of consumers who adopted during the crisis (i.e., the red area below the adoption curve on Fig. 21). To compute the number of consumers who postponed adoption because of the crisis, we predict post-crisis adoption had there not been a crisis using our regression.[25] Under these assumptions, we predict that absent the crisis, 933.1483937 consumers would have adopted real-time pricing during the period it occurred. Of those, 427 (or 45.8%) adopted despite the crisis and at most 61.3960564 (or 6.6%) postponed adoption. The rest chose to forego adoption altogether. These results suggest that strategically timing adoption is not very prevalent among consuumers. Rather, the results from Table 4 coupled with our simple counterfactual analysis suggest consumers make one-shot decisions on whether to adopt real-time pricing and forego adoption when spot prices have been high prior to switching contract.

[23]We assume that consumers who postpone adoption wait until prices decrease and not longer.

[24]More precisely, we regress the weekly number of RTP adopters on week-of-year-on-region fixed effects and then predict the number of adopters absent a crisis by setting the corresponding fixed effects to the average value in previous years.

[25]We define a post-crisis period for which we set the corresponding fixed effects to their average pre-crisis values to predict post-crisis adoption in the absence of a crisis.

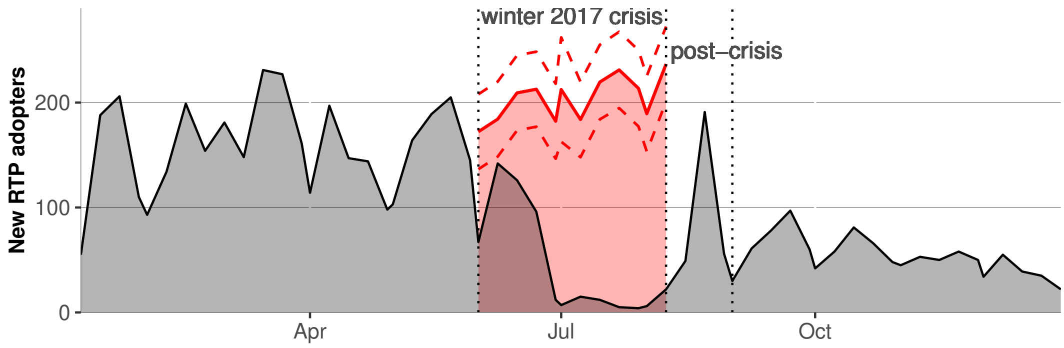


Figure 21 – Actual and predicted number of consumers joining RTP during and right after the crisis.

External validity

Note that our results come from a distinct subset of consumers. However, the existence of present bias likely extends to the other consumers as well. Consumers switching retailers, particularly those adopting real-time pricing, are likely to be more sophisticated than the rest of the population. Any bias found when analyzing sophisticated consumers is hence likely to exist among the general population as well. To see this, consider Table 1 in the previous section which summarizes the average observable characteristics of households in Wellington, among those switching retailers in the second half of 2014 and 2016, and among those adopting real-time pricing in the second half of 2014 and of 2016. On average, households switching retailers have higher incomes and are more educated than other households. Moreover, those adopting real-time pricing have significantly higher income and are more educated than any other consumer group.

In Table 8 in Appendix 3.B, we provide some evidence of how different demographic characteristics interact with price sensitivity. We use both *Recent Price* and *Future Price*, computed as last year's price and over 12 months, and interact both with income and education. In the first column, we see that a higher income makes consumers focus on future prices more, all else equal. In the second column, we see that consumers with more education react both less strongly to current prices and more to future prices. When combining all interaction terms into one regression in the third column, we see that, a bit surprisingly, higher-income individuals tend to put relatively more weight on current spot prices. However, this effect is dominated by the effect of education, which goes in the expected directions.

3.5 Discussion

Our results from Section 3.3 suggest that attrition during the winter 2017 crisis was higher among inexperienced consumers. Even though we were able to rule out that selection on (observable or unobservable) characteristics play a role, there are several reasons for why experience affects the decision to leave RTP during the crisis.

One explanation is that by experiencing real-time pricing, consumers can compare it to other tariffs that they know. Because spot prices are seasonal and volatile, it takes time to experience different situations and become familiar with real-time pricing. Therefore, experience allows consumers to form expectations about the long-term payoffs that they can derive from real-time pricing. These beliefs become more stable the more experience they get. As the comparison of tariffs on Figure 13 suggests, many consumers would likely have benefited from real-time pricing before the winter 2017 crisis started. Therefore, while all consumers on real-time pricing may have had optimistic beliefs about their long-term payoffs, the beliefs of experienced consumers were more entrenched. The winter 2017 crisis was a negative signal about the payoffs consumers can expect from real-time pricing. This result can explain why more inexperienced consumers than experienced ones felt the urge to react - by switching tariffs or lowering consumption.

Another possible explanation is that savings accumulated while on RTP have played a role. In fact, consumers on real-time pricing could check their weekly and cumulated savings on their mobile app and their online personal accounts. The app computes savings by comparing realized bills with the hypothetical bill consumers would have had to pay if they had stayed with their previous retailer and consumed identically. Figure 22 in Appendix 3.B provides two examples of how the app displayed the information to consumers during the winter 2017 crisis. The consumer from the left panel had adopted RTP several months before the crisis and, by the beginning of June 2017, had cumulated more than 1500$ savings. The consumer from the right panel had adopted RTP just as the crisis started, and after three weeks, her savings had always been negative. Due to a lack of granular consumption data, we are not able to test whether this mechanism is indeed prevalent. For similar reasons, we cannot study demand response of consumers and whether demand response was affected by the time consumers spent on RTp prior to the crisis.

The findings from Section 3.4 suggest that consumers who consider adopting real-time pricing make "now-or-never" decisions rather than strategically time adoption. This behavior is consistent with the fact that consumers often sign long-term contracts with their electricity retailer and, therefore, only have rare switching opportunities. In addition, we find that consumers strongly react to recent or ongoing spot prices. However, because spot prices are seasonal and volatile, ongoing spot prices are a poor predictor for long-term payoffs. Together, these observations suggest that adoption rates may be inefficiently low in periods of high price volatility because it is then likely that a consumer considering adoption faces high prices. Moreover, prospective adopters' reaction to ongoing spot prices suggest consumers may be present biased. This over-reaction of inexperienced consumers to short-run spot price fluctuations is also relevant to our findings about attrition: On the one hand, it concurs with the finding that recent adopter of real-time pricing are more likely to leave real-time pricing during the crisis and less likely to return after. On the other hand, it suggests that the behavior observed during the crisis may be an over-reaction.

3.6 Policy recommendations

We have presented evidence that consumers who are inexperienced with real-time pricing strongly react to recent or ongoing spot prices but that this present bias diminishes as they become more familiar with the tariff. This finding is essential to help understand why the adoption of real-time pricing remains low in New Zealand and to conceive solutions to overcome barriers to real-time pricing there and elsewhere.

As mentioned in the introduction, economic theory predicts that after introducing real-time pricing a retail electricity market would unravel to a situation with a large share of consumers on real-time pricing. However, unraveling did not occur in New Zealand, where less than 1.25% of households were on real-time pricing more than seven years after its introduction.

We have provided evidence that consumers who are inexperienced with real-time pricing - prospective and recent adopters - are highly sensitive to ongoing spot prices. While spot prices were fairly stable during the two years preceding the winter 2017 crisis, there were important price spikes[26] in winter 2017 and prices remained so afterward. This increase in spot price variations coincides with the sudden interruption in the growth of real-time pricing. The combination of important and unexpected spot price variations and consumer present bias could have jammed the unraveling process because of insufficient adoption and a reversal of the unravelling process during the crisis when many consumers left real-time pricing. Because countries plan to produce large amounts of electricity from intermittent renewable electricity, spot price volatility may increase. Similarly, spot price surges due to extreme weather events, such as in Texas in early 2021, may also become more common due to climate change, making the issues outlined in this paper even more salient. Therefore, it is relevant to think about

[26]We refer to weekly or monthly spot price variations rather than intra-day volatility.

ways to address them. We derive two sets of recommendations based on our findings: timing of adoption and information provision.

Essentially, present bias and lack of experience generate a setting where retailers or policy makers introducing real-time pricing need to be "lucky" and hope that no unexpected period of high price spikes arises during which large amounts of consumers leave real-time pricing and jam or revert an ongoing unraveling process. However, more experienced consumers seem to be less likely to leave during such events, strategically timing the introduction of real-time pricing may help prevent the take-up of real-time pricing to stall early on. Note that timing matters even in an opt-out set-up where consumers are defaulted into real-time pricing, as in Spain (Fabra et al., 2021), because of the risk of early attrition. However, timing the adoption process will not lead to wide-spread adoption in itself, especially when considering that in New Zealand, spot prices were low and stable during a full three years following the introduction of real-time pricing.

Our findings suggest that consumers use inadequate information to decide whether to adopt and then whether to discard real-time pricing. This may be either because computing long-term payoffs is costly or because they do not understand the spot price formation process. Either way, it is important to provide consumers with relevant and reliable information to help them make rational and informed decisions. First, consumers must be familiar with the formation of spot prices, the determinants of demand and supply such as the weather, and in particular, their seasonal patterns. The prevents consumers from underestimating the benefits of real-time pricing when spot prices are temporarily high. In addition, consumers need to be aware of whether, in the long run, they would benefit or not from real-time pricing. For that purpose, a simple policy would be to facilitate access to records of household consumption profiles and use them on tariff comparison websites. Ito et al. (2021) show that providing this information ex-ante to consumers significantly affects their choices by helping the structural winners self-select to time-varying tariffs.[27]

3.7 Conclusion

In this paper, we document the adoption of a new electricity tariff, real-time pricing, by residential consumers in New Zealand. Contrary to theoretical predictions, the retail market did not unravel and, more than seven years after the introduction of real-time pricing, less than 1.25% of consumers switched to this tariff. It is all the more puzzling that adoption

[27]Structural winners are consumers who benefit from adopting real-time pricing even without adjusting their electricity consumption

appears financially beneficial. We find that consumers inexperienced with real-time pricing - prospective and recent adopters - are highly sensitive to ongoing spot prices. The combination of present bias and lack of experience and unexpected periods of high spot prices may explain the puzzle: price spikes lead to low adoption and large attrition if they were not expected. However, with experience, consumers on real-time pricing focus less on ongoing events. Based on these findings, we make recommendations to help overcome these barriers to real-time pricing. We have derived two types of recommendations: timing of adoption - to increase adoption and limit attrition - and information provision - to help consumers make informed and rational decisions.

This paper opens several promising alleys for future research. First, as discussed in Section 3.6, because consumers are present biased, eliciting the optimal timing for advertising campaigns to encourage adoption in an opt-in set-up and for defaulting consumers to real-time pricing in an opt-out set-up is an open question. Second, because consumer learning about real-time pricing is important and there may be market failures associated with it, it may be efficient to subsidize the adoption of real-time pricing (in an opt-in set-up). Typically, if there are agency costs such as risk aversion (e.g., liquidity constraints) or externalities (e.g., social learning), then adoption rates will be inefficiency low. Then, for instance, consumers could receive a transfer after they have remained on real-time pricing a certain amount of time. The transfer should not depend on the consumption while on the tariff to avoid creating distortions.[28] A retailer incurring the costs of experimentation may not be able to recover them ex-post if, once the consumer has learned her valuation, she can switch to a competitor offering a low price. Thus, public subsidies may be required, and finding the optimal scheme is an open research question.

[28]In the case of New Zealand, the retailer Flick Electric guarantees its new customers that after 12-month they will have positive savings. This guarantee risks creating perverse incentives to over-consume during a crisis on the wholesale market.

3.A Appendix

3.B Additional tables and figures

Table 5 – Switching back to RTP after the winter 2017 crisis.

	Dependent variable:					
	Return to RTP (vs. Not)			Leave and Return (vs. Stay)		
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (month)	0.076***	0.077***	0.048***	0.027***	0.027***	0.018**
	(0.008)	(0.008)	(0.009)	(0.007)	(0.007)	(0.008)
Joined Last			−1.017***			−0.426***
			(0.159)			(0.156)
Winter consumption (MWh)		0.032	0.038		0.213***	0.217***
		(0.091)	(0.092)		(0.081)	(0.081)
log(Income) (k$/yr)		0.275	0.245		0.299*	0.287*
		(0.174)	(0.175)		(0.159)	(0.159)
Age		0.005	0.006		0.003	0.003
		(0.007)	(0.007)		(0.006)	(0.006)
Work (%)		0.002	0.001		−0.004	−0.004
		(0.006)	(0.006)		(0.005)	(0.005)
Education (%)		0.001	0.002		−0.003	−0.003
		(0.006)	(0.006)		(0.006)	(0.006)
Location-on-Previous retailer FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,494	2,378	2,378	12,925	12,265	12,265
Log Likelihood	−1,352.011	−1,281.669	−1,259.382	−2,277.045	−2,152.770	−2,148.813
Akaike Inf. Crit.	2,786.021	2,655.338	2,612.764	4,646.090	4,407.539	4,401.627
Note:						*p<0.1; **p<0.05; ***p<0.01

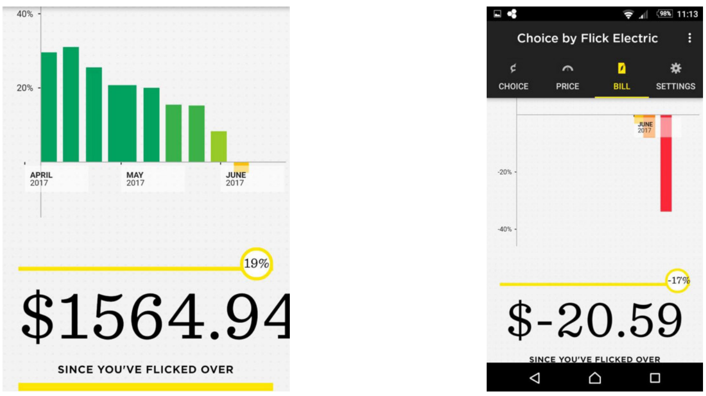


Figure 22 – Screenshots (obtained online) of the display of two customers' cumulative savings on their mobile application.

Table 6 – Correlation between recent price and future expected price.

		Future price			
		1 month	3 months	6 months	12 months
Recent price	2 weeks	0.71	0.40	0.31	0.08
	1 month	0.66	0.32	0.28	0.02

Note:
On a given day d, the recent price is computed as the average spot price in the period preceeding and ending at day d. The future price is computed as the average spot price in the period following and starting at day d+1.

Table 7 – Logit regression of switchers to non-traditional tariffs. We use data from 2014-09-01 to 2018-06-01

	Dependent variable:		
	Individual decision to adopt RTP		
	1 Week	2 Weeks	4 Weeks
	(1)	(2)	(3)
Recent Price	-0.084^{***}	-0.109^{***}	-0.155^{***}
	(0.008)	(0.010)	(0.011)
Volatilty	0.005	-0.006	0.011
	(0.007)	(0.011)	(0.017)
Consumption (MWh)	0.036^{***}	0.037^{***}	0.037^{***}
	(0.005)	(0.005)	(0.005)
Seasonal Difference (MWh)	0.015	0.015	0.015
	(0.052)	(0.052)	(0.052)
Income	0.004^{***}	0.004^{***}	0.004^{***}
	(0.001)	(0.001)	(0.001)
Age	0.006^{***}	0.006^{***}	0.006^{***}
	(0.002)	(0.002)	(0.002)
White collar worker	0.619^{***}	0.617^{***}	0.619^{***}
	(0.147)	(0.147)	(0.148)
Education	1.350^{***}	1.335^{***}	1.342^{***}
	(0.161)	(0.161)	(0.161)
Winter Crisis	-1.880^{***}	-1.676^{***}	-1.275^{***}
	(0.099)	(0.105)	(0.115)
Post Crisis	-0.979^{***}	-0.872^{***}	-0.679^{***}
	(0.079)	(0.080)	(0.084)
Location FE?	Yes	Yes	Yes
Month FE?	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes
Observations	34,000	34,000	34,000
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 8 – Logit regression of switchers to non-traditional tariffs with interaction effects between Recent/Future prices and demographics. We use data from 2014-09-01 to 2018-06-01

	Dependent variable:		
	Individual decision to adopt RTP		
	12 Month	12 Month	12 Month
	(1)	(2)	(3)
Recent Price	-0.159^{***}	-0.235^{***}	-0.169^{***}
	(0.020)	(0.015)	(0.020)
Future Price (Last Year)	-0.046^{***}	-0.058^{***}	-0.038^{**}
	(0.017)	(0.014)	(0.017)
Recent Price x Income	-0.0003		-0.001^{***}
	(0.0002)		(0.0003)
Future Price (Last Year) x Income	-0.001^{***}		-0.0004^{**}
	(0.0001)		(0.0002)
Recent Price x Education		0.002^{***}	0.004^{***}
		(0.0004)	(0.001)
Future Price (Last Year) x Education		-0.003^{***}	-0.003^{***}
		(0.0003)	(0.0004)
Consumption (MWh)	0.036^{***}	0.036^{***}	0.036^{***}
	(0.005)	(0.005)	(0.005)
Seasonal Difference (MWh)	0.014	0.017	0.018
	(0.053)	(0.053)	(0.053)
Income	0.016^{***}	0.003^{***}	0.015^{***}
	(0.002)	(0.001)	(0.002)
Age	0.006^{***}	0.006^{***}	0.006^{***}
	(0.002)	(0.002)	(0.002)
White collar worker	0.639^{***}	0.646^{***}	0.658^{***}
	(0.148)	(0.149)	(0.148)
Education	1.302^{***}	2.897^{***}	1.496^{***}
	(0.162)	(0.408)	(0.483)
Winter Crisis	-0.654^{***}	-0.673^{***}	-0.670^{***}
	(0.121)	(0.121)	(0.121)
Post Crisis	-0.267^{***}	-0.292^{***}	-0.291^{***}
	(0.091)	(0.091)	(0.091)
Location FE?	Yes	Yes	Yes
Month FE?	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes
Observations	34,000	34,000	34,000
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

3.C Attrition - robustness checks

3.C.1 Alternative measures of electricity consumption

Table 9 – Discarding real-time pricing during the winter 2017 crisis.

	Dependent variable:					
	Discard RTP					
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP(month)	-0.022^{***}	-0.021^{***}	-0.022^{***}			
	(0.005)	(0.005)	(0.005)			
log(Time on RTP)				-0.150^{***}	-0.140^{***}	-0.140^{***}
				(0.037)	(0.038)	(0.039)
Joined Last	0.442^{***}	0.455^{***}	0.451^{***}	0.272^{**}	0.305^{***}	0.301^{***}
	(0.081)	(0.082)	(0.083)	(0.111)	(0.114)	(0.115)
Annual Consumption (MWh)	0.003			0.003		
	(0.006)			(0.006)		
Winter Consumption (MWh)		0.112^{**}			0.109^{**}	
		(0.054)			(0.054)	
Seasonal Difference (MWh)			0.318^{***}			0.315^{***}
			(0.069)			(0.069)
log(Income) (k$/yr)	0.079	0.079	0.078	0.081	0.079	0.079
	(0.100)	(0.100)	(0.100)	(0.100)	(0.101)	(0.100)
Age	-0.010^{**}	-0.010^{**}	-0.011^{**}	-0.010^{**}	-0.010^{**}	-0.011^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Work (%)	-0.005^{*}	-0.005^{*}	-0.006^{*}	-0.006^{*}	-0.006^{*}	-0.006^{*}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education (%)	−0.003	−0.003	−0.003	−0.003	−0.003	−0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Location-on-Previous retailer FE?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,648	8,394	8,329	8,648	8,394	8,329
Log Likelihood	−4,111.338	−3,991.336	−3,948.746	−4,112.781	−3,993.355	−3,950.847
Akaike Inf. Crit.	8,322.676	8,080.673	7,995.491	8,325.562	8,084.709	7,999.693
Note:						*p<0.1; **p<0.05; ***p<0.01

3.C.2 Consumers who leave and come back

Table 10 – Discarding real-time pricing during the winter 2017 crisis. Sample without consumers who leave RTP during the crisis and come back afterwards.

	Dependent variable:			
	Discard RTP			
	(1)	(2)	(3)	(4)
Time on RTP(month)	-0.039^{***}		-0.038^{***}	
	(0.006)		(0.006)	
log(Time on RTP)		-0.259^{***}		-0.246^{***}
		(0.042)		(0.042)
Joined Last	0.636^{***}	0.345^{***}	0.622^{***}	0.352^{***}
	(0.091)	(0.126)	(0.091)	(0.127)
Winter Consumption (MWh)	0.106^{*}	0.101	0.111^{*}	0.106^{*}
	(0.063)	(0.063)	(0.063)	(0.063)
log(Income) (k$/yr)	−0.010	−0.008	0.009	0.009
	(0.115)	(0.115)	(0.115)	(0.115)
Age	-0.011^{**}	-0.012^{**}	-0.011^{**}	-0.012^{**}
	(0.005)	(0.005)	(0.005)	(0.005)
Work (%)	-0.007^{*}	-0.007^{*}	-0.007^{*}	-0.007^{*}
	(0.004)	(0.004)	(0.004)	(0.004)
Education (%)	−0.003	−0.003	−0.002	−0.003
	(0.004)	(0.004)	(0.004)	(0.004)
Location-on-Previous retailer FE?	Yes	Yes	Yes	Yes
Observations	8,394	8,394	7,923	7,923
Log Likelihood	−3,238.409	−3,240.725	−3,161.318	−3,164.431
Akaike Inf. Crit.	6,574.819	6,579.449	6,420.637	6,426.861
Note:				*p<0.1; **p<0.05; ***p<0.01

3.C.3 Comparing Wellington and Auckland

Table 11 – Robustness check: Probability of discarding RTP (Location refers to Christchurch in columns 1-3 and Auckland in columns 4-6)

	Dependent variable:					
	Discard RTP					
	Wel-Chr	Wel-Chr	Wel-Chr	Auc-Chr	Auc-Chr	Auc-Chr
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (month)	-0.05^{***}	-0.04^{***}	-0.05^{***}	-0.05^{**}	-0.04^{*}	-0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Location	0.15	0.12	0.24	0.13	0.10	0.46
	(0.14)	(0.14)	(0.34)	(0.20)	(0.20)	(0.55)
Time on RTP x Location	-0.0005	0.01	0.02	-0.001	0.005	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Annual elec cons (MWh)	-0.02^{**}	-0.02^{**}	-0.02^{**}	-0.02^{**}	-0.02^{**}	-0.02^{**}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Win/Sum cons diff (MWh)	0.38^{***}	0.38^{***}	0.38^{***}	0.46^{***}	0.45^{***}	0.45^{***}
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
Income (k$/yr)	0.001	0.001	0.001	-0.001	-0.0004	-0.0004
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Age	-0.01^{*}	-0.01^{**}	-0.01^{*}	-0.01	-0.01	-0.01
	(0.005)	(0.005)	(0.005)	(0.01)	(0.01)	(0.01)
Work (%)	-0.01^{*}	-0.01^{*}	-0.01^{*}	-0.004	-0.004	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Education (%)	-0.002	-0.002	-0.002	-0.005	-0.01	-0.01
	(0.004)	(0.004)	(0.004)	(0.01)	(0.01)	(0.01)
Month FE?	No	Yes	No	No	Yes	No
Month-on-NRR FE?	No	No	Yes	No	No	Yes
Observations	6,104	6,104	6,104	4,923	4,923	4,923
Log Likelihood	$-3{,}086.96$	$-3{,}076.09$	$-3{,}072.24$	$-2{,}575.69$	$-2{,}563.30$	$-2{,}558.82$
Akaike Inf. Crit.	6,193.92	6,194.19	6,208.47	5,171.37	5,168.60	5,181.64

Note: *p<0.1; **p<0.05; ***p<0.01

3.D Adoption - robustness checks

Table 12 – Logit regression of switchers to non-traditional tariffs. Recent price always computed over 2 weeks. We use data from 2014-09-01 to 2018-06-01

	Dependent variable:								
	Individual decision to adopt RTP								
	1 Month	1 Month	1 Month	6 Months	6 Months	6 Months	12 Months	12 Months	12 Months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Recent Price	−0.146***	−0.118***	−0.113***	−0.104***	−0.123***	−0.114***	−0.120***	−0.157***	−0.114***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Future Price (realized)	0.053***			0.105***			0.434***		
	(0.010)			(0.016)			(0.023)		
Future Price (last year)		0.045			−0.094***			−0.147***	
		(0.029)			(0.010)			(0.012)	
Future Price (AR (1))			0.016**			0.047**			0.093**
			(0.007)			(0.024)			(0.047)
Consumption (MWh)	0.036***	0.036***	0.036***	0.035***	0.036***	0.036***	0.035***	0.036***	0.036***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Seasonal Difference (MWh)	0.017	0.017	0.016	0.019	0.015	0.016	0.017	0.013	0.016
	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
Income	0.004***	0.004***	0.003***	0.004***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.006***	0.006***	0.006***	0.007***	0.006***	0.006***	0.006***	0.006***	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
White collar worker	0.631***	0.628***	0.629***	0.634***	0.629***	0.630***	0.578***	0.619***	0.630***
	(0.149)	(0.149)	(0.149)	(0.149)	(0.149)	(0.149)	(0.150)	(0.149)	(0.149)
Education	1.317***	1.321***	1.318***	1.321***	1.301***	1.316***	1.376***	1.314***	1.316***
	(0.162)	(0.162)	(0.162)	(0.162)	(0.162)	(0.162)	(0.163)	(0.162)	(0.162)
Winter Crisis	−1.784***	−1.666***	−1.723***	−1.896***	−1.103***	−1.714***	−1.243***	−0.914***	−1.713***
	(0.102)	(0.099)	(0.103)	(0.106)	(0.116)	(0.102)	(0.103)	(0.116)	(0.102)
Post Crisis	−0.881***	−0.798***	−0.827***	−0.898***	−0.177*	−0.824***	−0.736***	−0.292***	−0.824***
	(0.081)	(0.079)	(0.081)	(0.081)	(0.104)	(0.080)	(0.081)	(0.090)	(0.080)
Location FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,677	33,677	33,677	33,677	33,677	33,677	33,677	33,677	33,677

Note: *p<0.1; **p<0.05; ***p<0.01

Table 13 – Logit regression of switchers to non-traditional tariffs. Recent price always computed over 4 weeks. All prices in logs. We use data from 2014-09-01 to 2018-06-01

	Dependent variable:								
	Individual decision to adopt RTP								
	1 Month	1 Month	1 Month	6 Months	6 Months	6 Months	12 Months	12 Months	12 Months
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(Recent Price)	-1.331^{***}	-1.159^{***}	-1.149^{***}	-1.046^{***}	-1.083^{***}	-1.136^{***}	-0.982^{***}	-1.360^{***}	-1.134^{***}
	(0.091)	(0.079)	(0.079)	(0.081)	(0.079)	(0.079)	(0.079)	(0.081)	(0.079)
log(Future Price (realized))	0.293^{***}			0.748^{***}			2.800^{***}		
	(0.076)			(0.135)			(0.182)		
log(Future Price (last year))		0.230			-0.738^{***}			-1.633^{***}	
		(0.201)			(0.115)			(0.141)	
log(Future Price (AR (1)))			0.041			0.379^{**}			0.806^{**}
			(0.060)			(0.176)			(0.328)
Consumption (MWh)	0.035^{***}	0.035^{***}	0.035^{***}	0.035^{***}	0.035^{***}	0.036^{***}	0.035^{***}	0.035^{***}	0.036^{***}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Seasonal Difference (MWh)	0.020	0.019	0.018	0.022	0.018	0.018	0.019	0.017	0.018
	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
Income	0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	0.006^{***}	0.006^{***}	0.006^{***}	0.007^{***}	0.006^{***}	0.006^{***}	0.006^{***}	0.006^{***}	0.006^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
White collar worker	0.633^{***}	0.628^{***}	0.629^{***}	0.637^{***}	0.633^{***}	0.628^{***}	0.577^{***}	0.615^{***}	0.629^{***}
	(0.150)	(0.149)	(0.149)	(0.149)	(0.149)	(0.149)	(0.150)	(0.150)	(0.149)
Education	1.328^{***}	1.329^{***}	1.325^{***}	1.324^{***}	1.314^{***}	1.325^{***}	1.388^{***}	1.340^{***}	1.325^{***}
	(0.163)	(0.162)	(0.162)	(0.162)	(0.163)	(0.162)	(0.163)	(0.163)	(0.162)
Winter Crisis	-1.408^{***}	-1.347^{***}	-1.363^{***}	-1.602^{***}	-1.054^{***}	-1.394^{***}	-1.152^{***}	-0.764^{***}	-1.399^{***}
	(0.113)	(0.112)	(0.114)	(0.122)	(0.120)	(0.114)	(0.113)	(0.122)	(0.114)
Post Crisis	-0.612^{***}	-0.560^{***}	-0.575^{***}	-0.694^{***}	-0.192^{*}	-0.592^{***}	-0.452^{***}	-0.298^{***}	-0.593^{***}
	(0.087)	(0.086)	(0.088)	(0.090)	(0.104)	(0.087)	(0.087)	(0.090)	(0.087)
Location FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,677	33,677	33,677	33,677	33,677	33,677	33,677	33,677	33,677
Note:									*p<0.1; **p<0.05; ***p<0.01

Bibliography

Agan, A. and S. Starr (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Anderson, S. T., R. Kellogg, and J. M. Sallee (2013). What do consumers believe about future gasoline prices? *Journal of Environmental Economics and Management 66*(3), 383–403.

Ball, I. (2019). Scoring strategic agents. *arXiv preprint arXiv:1909.01888*.

Bar-Isaac, H. and J. Deb (2014a). (good and bad) reputation for a servant of two masters. *American Economic Journal: Microeconomics 6*(4), 293–325.

Bar-Isaac, H. and J. Deb (2014b). What is a good reputation? career concerns with heterogeneous audiences. *International Journal of Industrial Organization 34*, 44–50.

Bartik, A. and S. Nelson (2019). Deleting a signal: Evidence from pre-employment credit checks. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-137).

Borenstein, S. (2005a). The long-run efficiency of real-time electricity pricing. *The Energy Journal 26*(3).

Borenstein, S. (2005b). Time-varying retail electricity prices: Theory and practice. *Electricity deregulation: choices and challenges 4*, 317–356.

Bouvard, M. and R. Levy (2018). Two-sided reputation in certification markets. *Management Science 64*(10), 4755–4774.

Busse, M. R., D. G. Pope, J. C. Pope, and J. Silva-Risso (2015). The psychological effect of weather on car purchases. *The Quarterly Journal of Economics 130*(1), 371–414.

Calzolari, G. and A. Pavan (2006). On the optimality of privacy in sequential contracting. *Journal of Economic theory 130*(1), 168–204.

Chade, H., J. Eeckhout, and L. Smith (2017). Sorting through search and matching models in economics. *Journal of Economic Literature 55*(2), 493–544.

Corbae, D. and A. Glover (2018). Employer credit checks: Poverty traps versus matching efficiency. Technical report, National Bureau of Economic Research.

Cortes, K. R., A. Glover, and M. Tasci (2020). The unintended consequences of employer credit check bans for labor markets. *Federal Reserve Bank of Kansas City Working Paper* (20-04).

Daglish, T. (2016). Consumer governance in electricity markets. *Energy Economics 56*, 326–337.

Electricity Authority (2018). 2017 Winter Review - Final report.

Fabra, N. (2021). The energy transition: An industrial economics perspective. *International Journal of Industrial Organization*, 102734.

Fabra, N., D. Rapson, M. Reguant, and J. Wang (2021). Estimating the Elasticity to Real Time Pricing: Evidence from the Spanish Electricity Market. *American Economic Association Papers and Proceedings forthcoming*.

Farrell, J. and P. Klemperer (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of industrial organization 3*, 1967–2072.

Finlay, K. (2008). Effect of employer access to criminal history data on the labor market outcomes of ex-offenders and non-offenders. Technical report, National Bureau of Economic Research.

Fowlie, M., C. Wolfram, C. A. Spurlock, A. Todd, P. Baylis, and P. Cappers (2020). Default effects and follow-on behavior: evidence from an electricity pricing program. NBER Working Papers 23553, National Bureau of Economic Research.

Frankel, A. and N. Kartik (2019). Muddled information. *Journal of Political Economy 127*(4), 1739–1776.

Gibbons, R. and K. J. Murphy (1992). Optimal incentive contracts in the presence of career concerns: Theory and evidence. *Journal of political Economy 100*(3), 468–505.

Grubb, M. D. and M. Osborne (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review 105*(1), 234–71.

Harbaugh, R. and E. Rasmusen (2018). Coarse grades: Informing the public by withholding information. *American Economic Journal: Microeconomics 10*(1), 210–35.

Hirshleifer, D., B. Lourie, T. G. Ruchti, and P. Truong (2020). First impression bias: Evidence from analyst forecasts. *Review of Finance*.

Hölmstrom, B. (1979). Moral hazard and observability. *The Bell journal of economics*, 74–91.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The review of Economic studies 66*(1), 169–182.

Horner, J. and N. S. Lambert (2016). Motivational ratings.

Hortaçsu, A., S. A. Madanizadeh, and S. L. Puller (2017). Power to choose? an analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy 9*(4), 192–226.

Ito, K., T. Ida, and M. Tanaka (2021). Selection on welfare gains: Experimental evidence from electricity plan choice. NBER Working Papers 28413, National Bureau of Economic Research.

Joskow, P. and J. Tirole (2006). Retail electricity competition. *The RAND Journal of Economics 37*(4), 799–815.

Joskow, P. L. and C. D. Wolfram (2012). Dynamic pricing of electricity. *American Economic Review 102*(3), 381–85.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Lamp, S. (2018). Sunspots that matter. Working Paper.

Lerner, J. and J. Tirole (2006). A model of forum shopping. *American economic review 96*(4), 1091–1113.

Malmendier, U. and L. S. Shen (2019). Scarred consumption. NBER Working Papers 24696, National Bureau of Economic Research.

Miravete, E. J. (2003). Choosing the wrong calling plan? Ignorance and learning. *American Economic Review 93*(1), 297–310.

Mukherjee, A. (2008). Career concerns, matching, and optimal disclosure policy. *International Economic Review 49*(4), 1211–1250.

Nöldeke, G. and L. Samuelson (2015). Investment and competitive matching. *Econometrica 83*(3), 835–896.

Oxford Institute for Energy Studies (2019). Liberalized retail electricity markets: What we have learned after two decades of experience?

Padilla, A. J. and M. Pagano (1997). Endogenous communication among lenders and entrepreneurial incentives. *The review of financial studies 10*(1), 205–236.

Poletti, S. and J. Wright (2020). Real-time pricing and imperfect competition in electricity markets. *The Journal of Industrial Economics 68*(1), 93–135.

Rodina, D. (2016). Information design and career concerns. *Northwestern University Job Market Paper*.

Tirole, J. (2021). Digital dystopia. *American Economic Review 111*(6), 2007–48.

Wolak, F. A. (2013). Economic and political constraints on the demand-side of electricity industry re-structuring processes. *Review of Economics and Institutions 4*(1), 42.

Wozniak, A. (2015). Discrimination and the effects of drug testing on black employment. *Review of Economics and Statistics 97*(3), 548–566.