NBER WORKING PAPER SERIES

PUSHING THE ENVELOPE:
THE EFFECTS OF SALARY NEGOTIATIONS

Zoë B. Cullen
Bobak Pakzad-Hurson
Ricardo Perez-Truglia

Working Paper 33903
http://www.nber.org/papers/w33903

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2025

We thank Hashir Baqai, Zaheer Mohiuddin, Zuhayeer Musa, and Brian Nguyen for their help, and we thank levels.fyi for granting access to their data. We gratefully acknowledge research support from the National Science Foundation, UC Berkeley's Center for Equity, Gender, and Leadership, and Brown's Bravo Center for Economic Research. The collaborating institution did not provide any financial support for the research being conducted. We are also thankful for feedback from colleagues and seminar audiences at Princeton, Stanford (SITE), University of Chicago, University of Pittsburgh (BEDI), London School of Economics, UCLA, University of Stockholm, Harvard (Opportunity Insights), Harvard Business School, University of Taiwan, Boston University, Brown, NHH, UC-Berkeley, Google, ASSA Annual Meetings, Utah Winter Business Economics Conference, and the International Conference on Gender and Economics. This project was reviewed and approved in advance by the Institutional Review Board at Harvard Business School. The field experiment was pre-registered in the AEA RCT Registry (#0011491). Yuerong Zhuang, Julia Gilman, Miriam Malament and Alexia Witthaus Vine provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Pushing the Envelope: The Effects of Salary Negotiations
Zoë B. Cullen, Bobak Pakzad-Hurson, and Ricardo Perez-Truglia
NBER Working Paper No. 33903
June 2025
JEL No. C9, D80, J30, J38, J7

ABSTRACT

Salary negotiations are a widespread phenomenon that can shape key labor market outcomes, such as welfare and inequality. We provide novel empirical and theoretical insights into the causes and consequences of salary negotiations. We conducted two field experiments involving over 3,100 job seekers in the U.S. tech sector, designed to examine two types of information frictions. We find that a light-touch encouragement intervention significantly increased both negotiation attempts and compensation gains. In contrast, providing a substantial discount on negotiation coaching did not significantly affect negotiation attempts. Women responded more strongly to both interventions, helping to narrow gender gaps. We develop a new model of salary negotiations, incorporating risk and information frictions, that can better explain our experimental and non-experimental findings. The model's equilibrium analysis indicates that policies encouraging negotiation can enhance both welfare and equity.

Zoë B. Cullen
Harvard University
Harvard Business School
and NBER
zcullen@hbs.edu

Ricardo Perez-Truglia
University of California, Los Angeles
and NBER
ricardotruglia@gmail.com

Bobak Pakzad-Hurson
Brown University
bph@brown.edu

A randomized controlled trials registry entry is available at
https://www.socialscienceregistry.org/trials/11491

1 Introduction

Salary negotiations are a common feature of labor markets, with the potential to shape key outcomes such as compensation levels, pay inequality, and the gender wage gap. Yet despite their prevalence and potential importance, there is limited direct evidence on how salary negotiations unfold in practice and what effects they have. In this paper, we use survey and experimental data from active job seekers in the U.S. tech sector to uncover new empirical facts about the negotiation process. While existing workhorse models of negotiation under full information (e.g., Mortensen and Pissarides, 1994; Postel-Vinay and Robin, 2002; Cahuc et al., 2006) have notable merits, they cannot rationalize some key experimental and non-experimental findings. We therefore develop a new theoretical framework that accounts for the new facts and enables counterfactual and policy analysis.

A central empirical motivation for our study is the striking variation in negotiation behavior: some employees attempt to negotiate job offers, while others do not. Yet those who do negotiate are often successful, securing significantly better compensation terms. This pattern raises a central question: why don't all employees attempt to negotiate? The lack of negotiation is not necessarily inconsistent with complete information. For example, some employees may forgo negotiation because they do not have leverage and thus have no chance of improving their offer. However, based on anecdotal accounts, we hypothesize that employees' reluctance to negotiate may stem, at least in part, from information frictions.

We consider two types, or margins, of uncertainty. *Extensive-margin uncertainty* posits that employees may refrain from attempting to negotiate because they believe that the employer is not open to negotiations. For example, employees may have a fear—real or perceived—that the offer may be take-it-or-leave-it and thus attempting to negotiate would carry significant risks, such as a rescinded offer or a strained employer relationship. On the other hand, *intensive-margin uncertainty* posits that employees may refrain from negotiating because they feel like they lack the information and skills necessary to conduct the negotiation. For example, they may be unsure about how much to ask for, or how best to draft an email detailing their requests. These information frictions are noteworthy not only to the extent that they could explain why employees do not negotiate more often, but also because they can have welfare and policy implications. In a world with full information, employees who choose not to negotiate might be right not to—for instance, recognizing that without leverage, they have nothing

to gain. However, in the presence of information frictions, employees who forgo negotiation may leave money on the table, or may end up matching with an employer who does not value them most.

To shed light on how salary negotiations play out in real-world settings, we conducted a survey targeting professionals in the U.S. tech sector who were actively navigating the job market. Our research was conducted in partnership with levels.fyi, a company that provides detailed compensation data. We invited a large sample of their users to participate in our study. The baseline survey captured detailed information on participants' backgrounds and included an experiment with two treatment arms. Starting 45 days later, we conducted a follow up survey to track key outcomes, such as whether the subject attempted to negotiate a job offer and whether the offer improved from initial to final compensation terms.

The first treatment arm, called the *encouragement treatment*, was designed to address uncertainty at the extensive margin. This light-touch informational intervention encouraged participants to negotiate the terms of their job offers. In addition to a narrative promoting negotiation, the message included factual survey data on the share of individuals who attempt to negotiate and the proportion of those attempts that successfully result in higher compensation. Prior to receiving this information, we elicited participants' beliefs about these two statistics. This enables us to examine heterogeneity in treatment effects based on prior beliefs, consistent with standard practice in information-provision designs.

The second treatment arm, referred to as the *coaching treatment*, was designed to address uncertainty at the intensive margin. Levels.fyi offers a negotiation coaching service in which individuals are assigned to an experienced coach who provides personalized one-on-one guidance to equip candidates with the knowledge and skills needed to negotiate their job offers. The service is currently priced at $1,250 for the mid-level package and $2,450 for the senior package. In this treatment arm, details about the service were described to both the treatment and control groups. However, only the treatment group received a deep discount (80%+ off) to access the service. Moreover, before the offer-provision stage—and regardless of treatment assignment—we measured *stated* willingness to pay for the coaching service using an incentive-compatible method. These data allow us to explore heterogeneity between individuals with higher versus lower willingness to pay (WTP) for the service.

We recruited 3,858 subjects who completed both surveys. In the control group, the average participant earned $221,347 annually and had 7 years of work experience.[1] Common roles included

[1] Statistics refer to subjects in the control group, who did not receive either treatment.

Software Engineer, Product Manager, Data Scientist, and Director, with top employers such as Google, Meta, and Apple. Most respondents (73%) were employed but actively job searching. About 82% received at least one job offer; of those, 47.6% attempted to negotiate, and 48.2% of negotiators succeeded in increasing their compensation. Ultimately, 76.0% accepted one of the offers.

Some basic descriptive statistics suggest that the information frictions we hypothesize could be a plausible barrier to negotiation. Suggestive of extensive-margin uncertainty, when asked open-ended about the main barriers preventing people from negotiating, coaches cited fear of retaliation as a key factor. Also in open-ended questions, some participants explicitly mentioned concerns about backlash, and a few of them mentioned having already experienced it—such as job offers being withdrawn or relationships with employers being strained. These concerns may be especially salient in our context, where 40% of initial offers are made verbally, allowing employers to withhold written offers without formally retracting them. And suggestive of intensive-margin uncertainty, nearly half of participants had no prior experience negotiating salaries; and while some had read a book or taken a course, most reported no formal training in negotiation.

The results of the encouragement treatment arm indicate that extensive-margin uncertainty is a significant barrier to negotiations. The encouragement treatment significantly increased both the likelihood of negotiating and the probability of securing better compensation. More precisely, the treatment raised the rate of negotiation attempts by 7.3 percentage points (pp), from 53.7% to 61.0% ($p = 0.008$), and the share who experienced compensation gains by 6.5 pp, from 30.3% to 36.8% ($p = 0.012$). These effects imply that a great majority of the marginal negotiation attempts induced by the treatment were successful, suggesting that many individuals underestimated their ability to improve their offers. And as predicted by the belief-updating channel, the effects were concentrated among individuals whose prior beliefs underestimated how common or successful negotiations typically are.

By contrast, the results from the coaching treatment suggest that intensive-margin uncertainty is not a major barrier to negotiation. Demand for the service was very low—in the control group (who faced the full price) only 1.1% of subjects took up the service. The treatment had a statistically significant but economically small effect on take-up: even in the treatment group, which received an 80%+ discount, only 2.9% took up the service. The effects were somewhat stronger among individuals with higher stated WTP, but remained small even within that group. Reflecting the limited take-up, the coaching treatment had no meaningful effect on negotiation attempts, increasing the

likelihood of negotiating by an insignificant 0.9 pp ($p = 0.752$). Overall, limited access to affordable coaching does not appear to be a key factor explaining why individuals choose not to negotiate.

The above evidence indicates that among the general population of tech employees, intensive-margin frictions do not appear to be a major barrier to negotiation. However, there may be a minority of individuals for whom intensive-margin frictions are significant. To explore this possibility, we conducted another wave, referred to as Wave 2, in which we aimed to recruit subjects with the highest demand for the coaching service. We included an organic link to our survey on the levels.fyi website, right below an easy-to-see "Book Now" button, to sign up for the coaching service. We recruited 1,586 subjects in this second wave. We successfully attracted a sample with high demand for the service, as shown by both the stated WTP and the actual take-up: in the control group (who faced the full price), 12.2% of Wave 2 control group took up the coaching service, compared to just 1.1% in Wave 1. Due to the higher demand, in Wave 2 the coaching treatment had a large effect on take-up—raising it by 24 pp ($p < 0.001$). Despite the strong effect on take-up, the treatment still had no significant impact on negotiation attempts (1.4 pp, $p = 0.648$). In sum, the evidence suggests that, even among individuals with high demand for coaching, most who use the service would have negotiated on their own anyway.

Gender differences in negotiations are often cited as a contributing factor to the gender pay gap (e.g., Babcock and Laschever, 2003; Sandberg and Scovell, 2013). For this reason, we examine whether the treatment effects differ by gender. The tech sector may be particularly susceptible to gender biases, as it is a male-dominated industry, with women holding approximately 25% of technical roles in large U.S. tech companies (Deloitte, 2022). Indeed, there appear to be some relevant gender differences at baseline, with women receiving lower initial offers, being less likely to attempt negotiating, and consequently less likely to experience compensation improvements. We find suggestive evidence that both the encouragement and coaching treatments were more impactful for women than for men. The encouragement treatment was effective at increasing negotiations (and compensation gains) for both men and women, but about three times as strong for women as for men. For the coaching treatment, we find that men who took up the service would have negotiated on their own (and experienced compensation gains) anyways; by contrast, most women who took up the coaching service would not have negotiated on their own (or experienced negotiation gains). The results have to be taken with a grain of salt, as in both waves women are a small share of the subjects and thus the estimates for them are less precisely estimated. However, the consistent direction of the heterogeneity

across outcomes and treatments lends support to the existence of meaningful gender differences.

To assess whether our results are surprising or predictable, we conducted an expert forecast survey. We recruited 117 academic experts with relevant publications, described the experiment to them, and asked them to predict the treatment effects. The experts largely missed the mark. For instance, most predicted that the coaching treatment would be more effective than the encouragement treatment in increasing negotiation rates and compensation—whereas we find the opposite.[2]

Motivated by the evidence, we propose a theoretical model of salary negotiations which we use as a lens to interpret our empirical results and to explore welfare and policy implications. While existing workhorse models of labor market bargaining offer valuable insights, our model incorporates three key features that are jointly necessary to account for the patterns we observe. First, negotiation involves "pushing the envelope"—that is, employees are not passive price takers and can choose to counter employers' initial offers. Second, workers face extensive-margin uncertainty about offer negotiability—that is, they are unsure whether negotiating will lead to backlash, such as the offer being withdrawn. Third, we allow employees to receive multiple job offers, framing bargaining under risk as a portfolio choice problem à la Chade and Smith (2006).

Upon receiving job offers from a set of employers, an employee elects to negotiate by making counteroffers to any subset of them. Extensive-margin uncertainty may dissuade the employee from negotiating altogether. She weighs the risks and rewards of negotiating with one employer, conditional on her decision to bargain with others. For example, an employee may choose to negotiate with one employer only if she has a sufficiently compelling "safety" offer from another employer with which she does not negotiate. We characterize the employee's optimal bargaining strategy and show that it follows a "top down" structure: she initiates bargaining with the employers offering the highest match value and negotiates with additional employers only if the expected gains exceed the inside option provided by the rest of her portfolio in expectation.

The model helps rationalize our key experimental finding. Under standard complete-information models, an information treatment like our encouragement intervention should have no effect. In contrast, our model explains why the encouragement treatment increases the share of employees who attempt to negotiate and secure compensation gains. It also captures a series of non-experimental findings that conflict with predictions from canonical models. For example, standard bargaining

[2]Results reported in Appendix I.

models based on Bertrand competition (e.g., Postel-Vinay and Robin, 2002; Cahuc et al., 2006) predict that employees negotiate their lowest offers upward to match their best offer, leading to a reduction in the dispersion of final compensation terms. In sharp contrast, our model predicts that employees begin by negotiating their top offers, treating the others as fallback options, and leading to an increase in dispersion of compensation terms. Consistent with our model, the data show that employees are more likely to negotiate their best offers than their worst, and that negotiations increase the dispersion in final compensation terms.[3]

We leverage the model to examine welfare and policy implications, for which it is crucial to account for equilibrium effects. For instance, what would happen if our encouragement treatment were scaled across the entire market? In our experiment, we treat only a negligible fraction of employees, so it is reasonable to hold employers' wage-setting practices fixed. However, a market-wide rollout would likely lead to more widespread negotiation attempts, prompting employers to optimally adjust their wage-setting behavior in response (see Cullen and Pakzad-Hurson, 2023). To capture these dynamics, we embed the employee's negotiation problem into a richer framework that endogenizes employer responses. Employers choose both wage offers and whether to make them negotiable, aiming to maximize profits. In equilibrium, employees' beliefs about negotiability are consistent with employers' decisions.

A policymaker can use various levers to make salary negotiations more or less common in equilibrium. While numerous pay transparency policies have been enacted around the world (Perez-Truglia, 2023; Cullen, 2024), there are currently no policies specifically targeting salary negotiations. That said, several proposals have been put forward. Some aim to encourage *leaning-in*—that is, promoting negotiation—by offering free negotiation workshops (Enwemeka, 2016), launching information campaigns similar to our encouragement treatment, or imposing penalties on employers who retaliate against negotiation attempts. Other proposals aim to encourage *leaning-out*, for example by banning negotiations altogether (Kray, 2015). Existing models are largely silent on whether pushing in either direction is beneficial or harmful for welfare—and folk wisdom on the matter is divided.

Our model demonstrates that inducing a *lean-in* equilibrium—in which employers are always open to negotiating—increases employee de facto bargaining power and achieves the highest possible efficiency across equilibria. The efficiency gains arise in part because bargaining is frictionless,

[3]Offer matching may be the predominant form that bargaining takes in other markets—most notably, universities are known for adjusting pay to match competing offers in academia.

eliminating any "waste" from failed negotiations. Additionally, employers are compelled to concede the maximum surplus to employees, aligning employee incentives to maximize wages with those of a social planner who seeks to maximize efficiency. Indeed, we show that counteroffers need not be made in a "lean-in" equilibrium, because the threat of a counteroffer is so high that employers capitulate to employees by making high initial wage offers. We also show that, consistent with our experimental findings on gender gaps, the lean-in equilibrium promotes greater equity among employees with similar labor market prospects. We formalize how the lean-in equilibrium could be achieved through an information campaign that scales up our encouragement treatment. Finally, we show that the *lean-out* equilibrium leads to the lowest-employee-surplus outcome in certain labor markets.

Our paper contributes to several strands of literature. First, we contribute to the literature on bargaining in labor markets, offering both empirical and theoretical insights. On the empirical side, evidence on the role of salary negotiations remains limited. A seminal study by Hall and Krueger (2012) uses survey data to show that roughly one-third of U.S. employees engage in wage bargaining.[4] More recently, Caldwell et al. (2024) provides comprehensive survey evidence from Germany on the role of bargaining in labor markets. We contribute by uncovering novel evidence that information frictions and fear of employer backlash are key barriers to negotiation.

On the theoretical side, our model builds on existing frameworks while incorporating new features that we believe are essential for explaining key findings. Some standard approaches for bargaining are based on Nash's reduced-form bargaining solution (e.g., Mortensen and Pissarides, 1994), the alternating-offer model (e.g., Rubinstein, 1982; Stole and Zwiebel, 1996; Gentile Passaro et al., 2024), and Bertrand-like competition for employees (e.g., Postel-Vinay and Robin, 2002; Cahuc et al., 2006). More closely related to our approach, there are some bargaining models incorporating incomplete information (e.g., Myerson and Satterthwaite, 1983; Chatterjee and Samuelson, 1983; Fudenberg et al., 1985; Fuchs and Skrzypacz, 2010), although they capture something closer to intensive-margin uncertainty rather than extensive-margin uncertainty.[5] Although existing bargaining models offer valuable insights, they fail to capture the novel experimental and non-experimental findings that our model successfully explains.

Lastly, our study relates to and contributes to the literature on the gender pay gap. Salary ne-

[4] Salary negotiations are more common in high-skill positions (Hall and Krueger, 2012), presumably because these roles require more tailored offers (Cullen et al., 2022).

[5] In other words, the central tension agents must resolve is *how much* to ask for, not *whether* to ask for more.

gotiations are frequently cited as a contributing factor to gender disparities in pay (e.g., Card et al., 2016; Recalde and Vesterlund, 2023). Reflecting this concern, public discourse has increasingly encouraged women to "lean in" by negotiating more assertively (Babcock and Laschever, 2003; Sandberg and Scovell, 2013). The scientific evidence, however, is mixed. While women tend to negotiate significantly less in laboratory settings (e.g., Babcock and Laschever, 2003), doing so can lead to costly impasses (Exley et al., 2020; Dannals et al., 2021). Outside the lab, some survey studies find that the gender gap in negotiation may be less pronounced (Lachowska et al., 2022; Caldwell et al., 2024) and even reversed (Kray et al., 2024). Experimental studies with MBA students and online survey participants find that information provision can shift perceptions about the gender differences in negotiation, though the behavioral impacts are less clear (Capozza, 2024; Cortes et al., 2024). We contribute novel evidence from a field experiment conducted in a natural, high-stakes environment. In our context—a male-dominated industry—we find that both the encouragement and coaching interventions are relatively more beneficial for women, suggesting that such interventions could be leveraged to narrow the gender pay gap. Our findings are consistent with Roussille (2024), who studies the tech sector too, and shows that setting default asking salaries can reduce gender disparities.[6] More broadly, our theoretical model formally demonstrates that policies promoting a shift toward a lean-in equilibrium can enhance equity.

The remainder of the paper is organized as follows. Section 2 describes the institutional context, experimental design, and implementation details. Section 3 presents the effects of the encouragement treatment, while Section 4 covers the effects of the coaching treatment. Section 5 examines heterogeneity by gender. Section 6 introduces the model. The final section concludes.

2 Experimental Design

2.1 Overview of the Experimental Design

We collected panel data, consisting of a baseline survey and a follow-up survey. The main goal of the baseline survey was to administer the treatments. The main goal of the follow-up survey was to measure the outcomes of interest, such as whether the subject attempted to negotiate. A sample of the base-

[6]Also consistent with gender disparities in negotiations, evidence on public school teachers shows that the introduction of individual bargaining can widen the gender pay gap (Biasi and Sarsons, 2021).

line and follow-up survey instruments are attached as Appendices D and E, respectively. The survey was largely adaptive—for example, if an individual reported to have an offer, we would ask additional questions about that offer, like the compensation terms and whether he or she attempted to negotiate it. The full survey logic is detailed in Figure B.4. The surveys typically took between 5 and 10 minutes to complete.[7] To encourage participation in the study, we paid subjects $50 via electronic gift card.[8]

The baseline survey started with a series of questions that were used as filters. To qualify for the study, subjects must have: had permission to work in the U.S., been at least 20 years old, had a college degree, and either had an offer or been actively looking for one.[9] The survey collected identifiable information, including full name, email address, and employment history. In addition, we were able to match 78% of participants to a LinkedIn profile.[10] The LinkedIn data allows us to validate survey responses, capture additional subject characteristics, and track subsequent job mobility—even for participants who did not complete the follow-up survey.

The baseline survey gathered detailed subject characteristics, including demographics, prior negotiation experience or training, and information about the respondent's current or most recent job—such as employer, job title, and compensation. To standardize compensation reporting, we closely followed the approach used by levels.fyi. Specifically, we asked separately about annual base salary, annual equity compensation, annual bonuses, and relocation or sign-on bonuses. Unless otherwise noted, *compensation* refers to annual total cash compensation following the definition used by levels.fyi: the sum of annual base salary (on average, 77.9% of the total), annual equity compensation (15.9%), and annual bonuses (6.2%).[11] This definition implicitly assumes that employees value these three components roughly equally. A simple analysis of how individuals choose between offers provides strong support for that assumption—see Appendix C.2.

The goal of the follow-up survey was to track how job search unfolds—for example, whether participants receive offers and attempt to negotiate them. One implementation challenge was the

[7] The median completion times were 8.38 minutes for the baseline survey and 5.25 minutes for the follow-up survey.

[8] There were some small variations in these financial incentives—see Appendix B.1.

[9] For more details, see Appendix B.1. Additionally, we used data on IP addresses and LinkedIn profiles to confirm that the vast majority of subjects were indeed based in the U.S.

[10] In the first wave, we manually matched respondents to their LinkedIn profiles. Beginning in the second wave, we required participants to provide a valid LinkedIn profile at the start of the survey— see Appendix B.5.

[11] In practice, equity may vest unevenly over several years. Following the levels.fyi definition, we asked participants to report the average annual value across all vesting years. Since relocation and sign-on bonuses are one-time payments, following levels.fyi's definition we exclude them from the definition of annual compensation. In any case, these bonuses are relatively small—on average, only 3.5% relative to a year's worth of compensation.

varying pace of job searches across individuals, so we staggered the invitation emails to accommodate this variation. Approximately 45 days after the baseline survey, we sent the first round of follow-up invitations. If a participant responded but had not yet received any offers, we sent a second invitation 45 days later. If they still had not received any offers by that second checkpoint, we sent a final invitation four weeks later. Participants who did not respond to an invitation were sent multiple reminders. Among individuals who completed the follow-up survey, the median time between the baseline and follow-up surveys was 56 days.[12]

In the baseline survey, we asked participants whether they had any job offers at the time. For those who did, we collected detailed information, including compensation terms. To keep the baseline survey manageable, we asked for details on only one offer—if they had multiple, we asked them to report their preferred one.[13] In the follow-up survey, participants were reminded about the offer they listed at baseline, and they were asked to list any additional offers and to provide details for each. We allowed individuals to list up to three new offers, including any offers from their current employer. A small minority of participants (2%) had more than three additional offers; in those cases, we asked them to report their top three. For each offer reported (including the one from the baseline), the follow-up survey included a series of questions used to construct outcome variables, such as whether the participant attempted to negotiate and whether the compensation terms improved—more details in Section 2.4 below.

2.2 Encouragement Treatment Arm

The two treatment arms were located toward the end of the baseline survey. Treatment assignment was cross-randomized, so that 25% received the encouragement treatment, 25% received the coaching treatment, 25% received both, and 25% received neither.

The *encouragement treatment arm* targeted what we call extensive-margin uncertainty: people may choose not to negotiate because they believe the employer is not open to negotiations. This treatment consisted of a light-touch informational intervention that encouraged participants to negotiate job offers. Half of the subjects randomly assigned to the encouragement treatment saw an additional screen with the encouragement message (reproduced in Panel A of Figure 1), while

[12] For more details, see Appendices B.4 and C.1.

[13] Subjects with multiple offers at baseline were a minority: approximately 12.8% had two offers, and 6.4% had three or more.

the other half—the control group—did not see this screen. In addition to a narrative encouraging negotiations, the message included factual information from a recent survey (Fidelity, 2022) stating that 42% of individuals attempt to negotiate offers and that 85% of those who do are successful.[14] The message also included additional arguments in favor of negotiating—for instance, that one should not feel guilty about negotiating, that companies expect it, and a mention of a Harvard Business School course highlighting the long-term financial benefits of negotiating.

Given that the treatment message included information about the share of individuals who negotiate and their success rate, it is natural—under standard learning models—to expect heterogeneous effects by prior beliefs. For example, individuals who underestimate the share of negotiators should update their beliefs upward in response to the information and thus become more likely to negotiate themselves. By contrast, individuals with accurate beliefs should not update and therefore should not respond to the information. Indeed, leveraging heterogeneity by prior beliefs is a common approach in information-provision experiments (see e.g., Cullen and Perez-Truglia, 2022). With this in mind, we measured beliefs about the share of individuals who attempt to negotiate and their success rate immediately before the encouragement-provision stage. These data allow us to explore the (pre-registered) heterogeneity by prior beliefs.

2.3 Coaching Treatment Arm

The coaching treatment arm was designed to capture intensive-margin frictions: i.e., even if individuals are aware that negotiating is an option, some may refrain from negotiating simply because they feel they lack information or skills needed to succeed. Successful negotiations may involve multiple tasks. Individuals must determine what to ask for—should they negotiate for a higher base salary, a larger bonus, or more equity compensation? Or should they focus on non-compensation benefits instead? Even if they know what to ask for, pinpointing the right amount can be challenging, though websites like levels.fyi provide readily available data. Additionally, small tasks such as crafting negotiation emails could be intimidating. Even with a basic understanding, individuals may lack the confidence to negotiate and ultimately choose to forgo it altogether.

The coaching treatment consisted of a special offer providing a deep discount on a negotiation coaching service offered by levels.fyi. Before learning whether they received the discount, all

[14] To enhance credibility, the treatment message included hyperlinks to all the referenced articles.

subjects—both treatment and control—were shown a brief description of the service, including its main features and standard price. As part of the service, a subject was assigned to a dedicated coach who was available via video or phone for a three-month period.[15] During this time, the coach could assist with planning communications, drafting emails, and advising on outreach. The service was highly rated by users.[16] Reviewers praised the responsiveness and professionalism of the coaches and highlighted the personalized attention and tailored advice. Many reviews also mentioned that the sessions were effective not only for maximizing pay, but also for other purposes such as managing anxiety and making decisions about which offer to accept.

Half of the subjects, randomly assigned to the treatment group, were shown an additional screen (reproduced in Panel B of Figure 1) offering the service for a heavily discounted price ($250), while the other half—the control group—did not see this screen.[17] Individuals in the control group could still book the negotiation service through the levels.fyi website, but without the discount. In consultation with the partner organization, we purposely designed the treatment as a deep discount rather than making it entirely free.[18] We made it as easy as possible for recipients to claim the discount. They had the option to schedule their first meeting with the coach directly from the survey. If they did not do so, we later sent them a reminder via email informing them that the discount was still available.

There were two relevant packages—mid-level and senior-level—each tailored to candidates with different levels of experience and priced accordingly.[19] In our sample, about 59% of subjects fit into the mid-level service, while the remaining 41% fit in the senior-level.[20] The service cost $1,250 for the mid-level package and $2,450 for the senior package.[21] The implied discount was then 80% for the mid-level package (from $1,250 to $250) and about 90% for the senior package (from $2,450

[15] For more details, Figure B.2 shows a screenshot of levels.fyi's Frequently Asked Questions (FAQs) for the negotiation coaching service.

[16] Appendix B shows a screenshot from levels.fyi's website that lists customer reviews for the negotiation coaching service.

[17] In response to feedback from levels.fyi, we revised the terms of the offer during the initial months of recruitment, ultimately converging to these final terms—see Appendix B.3.

[18] One concern was that a free offer might be perceived as too good to be true. Another concern was that providing the service entirely for free could lead individuals to take it up without taking it seriously (e.g., not meeting with the coach).

[19] There was also a leadership-level package, which was tailored to a small minority and not relevant because it was not included in the experiment.

[20] In Wave 2, there were different URLs for individuals interested in the mid-level and senior-level services. In Wave 1, we used the subject's years of experience to determine whether they would be a better fit for the mid-level or senior-level package.

[21] Since recruitment for the two waves spanned nearly two years, there was one price change during the sample period: in February 2024, the company increased the price of the mid-level service from $650 to $1,250, and the price of the senior service from $1,887 to $2,450.

to $250).[22]

After introducing all subjects to the levels.fyi negotiation coaching service—but before informing them whether they would receive the discount—we elicited their WTP for the service. To incentivize truthful responses, we used an incentive-compatible method (Becker et al., 1963). More precisely, we employed a price-list method in which subjects made choices across five hypothetical scenarios. In the first scenario, they had to choose between the coaching service and a $100 Amazon gift card. In the remaining four scenarios, they faced the same binary choice but with increasing monetary amounts, up to $600. To introduce real stakes, we explained that 10 participants would be randomly selected to have one of their choices implemented—that is, there was a positive probability that one of their selections would have real consequences.[23] The willingness-to-pay data is valuable in its own right, as it provides insight into demand for the service. It also allows us to study heterogeneity by WTP. For example, individuals with a stated WTP above $250 should—at least in theory—be expected to take up the service when offered the discount, while those below that threshold should not take up the service even when offered the discount.

2.4 Outcomes of Interest

The first main outcome of interest is whether subjects attempted to negotiate. Specifically, the outcome variable equals 100 if the individual attempted to negotiate at least one offer and 0 otherwise.[24] As with any survey question, our measure of negotiation attempts may suffer from some measurement error.[25] The second main outcome is whether the compensation terms improved, regardless of whether a negotiation attempt occurred. Specifically, the outcome variable equals 100 if

[22]These calculations are based on the latest prices—the discounts were even deeper in percentage terms when using the prices from the earlier period.

[23]For the 10 subjects selected to have their choices implemented, the survey ended prematurely and they were not included in the subject pool.

[24]Some individuals may have received multiple offers. For these individuals, the measure captures the extensive margin—i.e., whether they attempted to negotiate any offer at all. There may also be intensive margin responses, such as negotiating a larger number of offers. However, this margin is limited: as discussed in Section 2.7, most individuals have a single offer and even those who do have multiple offers typically attempt to negotiate only one.

[25]For example, some individuals may not pay attention to the question, may misread it, or may make mistakes when responding. In addition, there is subjectivity in how respondents interpret what constitutes a negotiation attempt. For instance, delaying a decision on an offer can be a negotiation strategy—the employer may grow anxious and improve the offer proactively. As a result, some respondents may report that they did not attempt to negotiate, even though they may have done so implicitly by merely waiting. Indeed, we observe some suggestive evidence consistent with this: among offers for which the subject did not report a negotiation attempt, a small but non-negligible share (13.9%) experienced an improvement in compensation terms.

the annual compensation increased from the original terms to the latest terms for at least one offer, and 0 otherwise. Alternatively, we can measure this outcome as the percentage change in compensation terms from the initial offer to the final offer.[26]

There are two additional outcomes that are closely related to the negotiations outcome. First, we construct a measure that can pick up short-lived effects of the intervention. For instance, in the case of the encouragement treatment, individuals may have forgotten the information or it may have become less salient over time. At the end of the baseline survey—after subjects had been exposed to their assigned treatments—we asked whether they intended to negotiate in the future. We construct the intention to negotiate outcome that takes the value 100 if the subject expected to negotiate future offers almost surely, and 0 otherwise.[27] Second, we included two questions that can serve as a placebo outcome, for a falsification check. In the baseline survey, we asked subjects about whether they had experience negotiating offers in the past. Because this question was asked before subjects were exposed to any treatments, we should expect no treatment effects on this outcome. Likewise, we asked another question—also administered before treatment assignment—on whether they had attempted to negotiate when they received the offer for their current or most recent job. This constitutes our second placebo outcome.

In addition to the two main outcomes, the follow-up survey included a range of additional measures intended to capture the broader impact of the interventions or underlying mechanisms. For example, we asked individuals about their overall satisfaction with how their negotiations went, how well the offers matched their preferences, and whether they ended up changing employers (using either survey responses or tracking LinkedIn profiles). All these additional outcomes are described in detail in Appendix C.5.

2.5 Subject Recruitment

We designed our recruitment to target active job seekers in the United States—that is, individuals who had recently received a job offer or were currently interviewing and expecting to receive one in the

[26] To limit the potential influence of outliers, we winsorize this outcome at the 95th percentile. For individuals who experienced compensation improvements in multiple offers, we define the alternative outcome—to be consistent with the definitions of the other outcomes—as the maximum percentage change across all offers.

[27] We elicited this question using a 1 to 4 scale ranging from "No" to "Almost Surely." For ease of interpretation, the analysis focuses on a binary indicator. However, results are similar when using the full cardinal scale.

near future. In this section we describe the recruitment process for our main survey wave, which began on May 25, 2023, and ended on October 10, 2023. Individuals were required to register for a free account in order to access salary information on the levels.fyi website across different companies and positions. We sent email invitations to a sample of individuals who had recently registered an account.[28] The goal of this recruitment strategy was to obtain a representative sample of tech employees. Indeed, while the final sample is not perfectly representative of the tech sector, a comparison to a benchmark sample suggests that it is broadly similar across a wide range of characteristics (see Appendix C.6).

In the survey invitation (attached as Appendix H), we emphasized that the survey was intended for individuals actively participating in the job market.[29] Additionally, the registration form included information that allowed us to target invitations to individuals who reported having U.S. work authorization and being active on the job market.[30] The implied response rate was approximately 2.5% (see Appendix D). This response rate is somewhat lower than that of online surveys using similar methods (e.g., 4.7% on average, according to the meta-study by Sinclair et al. (2012)), but a lower response rate was expected given the context: many recipients were likely not actively on the job market, the compensation-related questions may have been too sensitive for some respondents, and given the volume of emails many of the invitations were likely filtered into spam folders.

In the first wave, a total of 2,435 subjects qualified for the study and completed the baseline survey.[31] Of those, 1,552 completed the follow-up survey, among whom 1,336 received at least one job offer. As shown in Appendix B, and consistent with successful random assignment, the pre-treatment characteristics are balanced across treatment and control groups.

2.6 Relevant Details about the Context

Our analysis focuses on compensation terms. This is of course not the only characteristic of the offer that matters to subjects, but it is definitively the most important. For example, 79% of

[28] Each week we invited individuals who had signed up on the platform during that same week. Additionally, we would invite a random sample of individuals who had signed up during the previous year. From August 2, 2023, to August 15, 2023, we also sent invitations to individuals who expressed interest in the coaching service via a form embedded on levels.fyi's website. A small minority (0.43%) of subjects were recruited this way.

[29] For each survey invitation, if an individual did not respond, we sent a reminder one week later and another reminder two weeks later.

[30] More precisely, we selected individuals who, when asked about their job market status, chose one of the following options: "actively interviewing," "actively searching," or "offer in hand".

[31] We consider a baseline survey complete if the respondent reached the stage at which they were randomized to see either a treatment or control screen.

participants accepted the offer with the highest total compensation.[32] However, subjects seem to value other terms of the offer too. A simple conditional logit model estimated using the offer choice data indicates that, while compensation is the most important factor, subjects also value other attributes such as remote work and the employer's reputation—see Appendix C.2. In any case, as discussed in Appendix C.5, changes in non-compensation terms are much rarer than changes in compensation.

In our baseline survey, we asked whether the job offer was verbal or written; approximately 40% were verbal. This is consistent with anecdotal reports that employers often extend verbal offers first and only formalize them in writing once they believe the candidate is prepared to accept. This practice matters for two reasons. First, verbal offers may make negotiation attempts riskier, as employers can more easily withdraw them with fewer legal consequences.[33] Second, verbal offers make it more difficult to match outside offers, insofar as employers must rely on the candidate's word rather than written proof.

When asked about the main barriers to negotiation, coaches frequently cited fear of retaliation as a key factor. In a small survey of coaches, we posed an open-ended question asking why some individuals choose not to negotiate. Of the 11 coaches who responded, 5 explicitly mentioned fear of backlash. We also asked participants an optional open-ended question about whether they were concerned about experiencing backlash from their negotiation attempts. Among the 36.5% who responded to this question, 9.9% mentioned concerns about potential backlash. While the risk may be low-probability, it is not zero: a small minority (1.9%) reported having experienced backlash already. For example, some respondents described having offers rescinded after attempting to negotiate, while others noted that employers reacted angrily to their negotiation attempts.

Another relevant feature of the context is the state of the job market during the study period. At that time, demand for software engineers was contracting. Hiring slowed significantly across employers, and there were widespread layoffs.[34] As a result, the returns to negotiations we estimate in this study are possibly lower than they would have been in a stronger job market. Indeed, the coaches largely agreed that negotiating offers had become more difficult compared to recent years.[35]

[32] More precisely, this is the share accepting the highest-paid offer among individuals with no current job and exactly two outside offers that differ by at least 5% in compensation.

[33] In the U.S., withdrawing a job offer is not automatically illegal because employment is usually at-will. However, the employer could still be sued, for example, if the candidate claims that he or she was discriminated against.

[34] This likely reflected a combination of two factors. First, in the two years following the pandemic, tech firms may have hired too aggressively and subsequently had to scale back. Second, the rapid growth of AI reduced the demand for software developers, particularly at the junior level (Business Insider, 2025).

[35] Among the 12 coaches who responded to our survey, 10 believed that negotiating offers was more difficult in

2.7 Descriptive Statistics

To provide a general sense of who the subjects are, we begin by presenting descriptive statistics for the control group—those who did not receive either treatment. The average subject was 31 years old and 20.0% were female. The average subject had 7 years of work experience, 51.1% had a college degree and 72.7% of them were currently employed at the time of the baseline. The average offer paid $217,039 in total annual compensation, which is typically significantly above their current or most recent compensation. Common job titles included Software Engineer, Product Manager, Data Scientist, and Director, with major employers such as Meta, Google, and Apple. About half of the subjects reported having prior experience with negotiations, and their self-rated confidence in negotiating was 6 on a scale from 1 to 10. Approximately 86.0% of these subjects received at least one offer.[36] Among them, 47.6% attempted to negotiate. Of those who negotiated, 48.2% experienced an improvement in compensation terms. Among the offers improved, the average increase in total compensation was 13.9%, or $34,817 per year. Ultimately, 76.0% accepted one of the offers.

3 Effects of the Encouragement Treatment

3.1 Distribution of Prior Beliefs

We begin by examining the distribution of prior beliefs, which is key to interpreting the effects of information. Figure 3 shows the distribution of prior beliefs, with Panel A corresponding to the perceived share of individuals who negotiate and Panel B corresponding to the share of those negotiations that are successful. There is large heterogeneity in prior beliefs. For instance, some subjects believed that nearly everyone negotiates, while other subjects believed that nearly nobody negotiates. This disagreement between subjects imply that *some* subject must be wrong. Most importantly for the interpretation of the effects of information, what matters the most is whether the information provided to subjects was above or below their prior beliefs, which determines the direction of the belief updating and therefore the direction of the expected effects.

In each panel of Figure 3, the feedback provided in the encouragement message is denoted with

2023–24 than in previous years, while only 1 believed that it was slightly easier.

[36] This can be broken down: 59.3% received exactly one offer and 18.0% received multiple offers (typically two).

a vertical line. A majority of subjects (62%) underestimated the share of negotiators, while a vast majority (89%) underestimated the success rate.[37] For simplicity, our baseline specification splits the sample into two groups. The *low-priors* group (roughly 70% of the sample) includes individuals whose prior beliefs were below the feedback for both questions, and who we would therefore expect to update their beliefs upward. The remaining 30% of subjects are in the *high-priors* group. In any case, since nearly all respondents underestimated the second prior belief, the split is driven almost entirely by whether individuals underestimated or overestimated the share of negotiators.

3.2 Average Treatment Effects

Figure 4 summarizes the effects of the encouragement treatment. Each panel corresponds to a different outcome and shows the average for the treatment and control groups, along with the difference—that is, the average treatment effect. For now, we focus on the left half of each panel, which presents the results for the full sample. The outcome variable in Panel A of Figure 4 is the ex-ante intention to negotiate, which was elicited toward the end of the baseline survey. In the control group, 77.4% of respondents reported that they would almost surely negotiate future offers. That fraction rises to 80.9% in the treatment group, indicating that the encouragement treatment had a positive and statistically significant effect of 3.6 pp ($p = 0.001$). In other words, the treatment made individuals more enthusiastic about negotiating—at least in the short term.

The effect on the intention to negotiate should be interpreted with caution due to the hypothetical nature of the question and the potential for experimenter demand effects. While individuals may express a greater willingness to negotiate, they may not follow through when faced with real decisions.[38] For this reason, the key outcome of interest is whether, months later, subjects had actually attempted to negotiate when given the chance. The results for this outcome are presented in Panel B of Figure 4. By construction, this outcome is defined only for individuals who responded to the follow-up survey and received at least one job offer—otherwise, they would not have had anything to negotiate.[39] In the control group, 53.7% of eligible subjects attempted to negotiate. The encouragement treatment increased

[37] For the analysis of the treatment effects, what matters is the comparison between the prior beliefs and the feedback provided. A different question is whether the prior beliefs are accurate or not—for a discussion, see Appendix B.2.

[38] Indeed, we observe that many individuals who stated at baseline that they would almost surely negotiate ultimately did not do so.

[39] Panel A of Figure 4 includes all individuals who completed the baseline survey. Panel A of Figure C.5 shows that the results from Panel A of Figure 4 are similar when applying the same sample restrictions.

this share by 7.3 pp (p = 0.008), from 53.7% in the control group to 61.0% in the treatment group.

Next, we examine the effects of the encouragement treatment on the probability of experiencing a compensation gain. Consider the marginal individuals who chose to negotiate as a result of the treatment. If their offer terms improved, this would suggest that not negotiating would have been a mistake. Conversely, if their negotiation attempts were largely unsuccessful, it would indicate that their initial reluctance to negotiate was justified.

Panel C of Figure 4 shows that the encouragement treatment increased the share of individuals with compensation improvements by 6.5 pp (p = 0.012), from 30.3% to 36.8%. This treatment effect on compensation improvements (5.7 pp) is nearly as large as the treatment effect on negotiation attempts (7.3 pp, from Panel B). Taken together, these results imply that 78.1% ($= \frac{5.7\%}{7.3\%}$) of the marginal negotiation attempts induced by the treatment were successful. In other words, the evidence suggests that for most individuals who were persuaded to negotiate by the encouragement treatment, not negotiating would have been a missed opportunity.[40]

3.3 Heterogeneity by Prior Beliefs

In each panel of Figure 4, the right side of the figure splits the sample into the low-priors and high-priors groups. Because the factual information about negotiation rates and success rates was not the *only* content of the information treatment, we cannot make unambiguous predictions about the direction of the effects for each group. However, the belief-updating mechanism yields a clear prediction about the relative magnitude of the effects: the treatment should have more positive effects for individuals in the low-priors group—who are updating both of their beliefs upward—than for those in the high-priors group, who are revising one or both beliefs downward.

Consistent with the belief-updating channel, Figure 4 shows that the effects are more positive for the low-priors group than for the high-priors group. In fact, the effects were largely concentrated in the low-priors group. Panel A shows that the treatment increased the intention to negotiate by 4.2 pp (p < 0.001) for the low-priors group, compared to just 2.2 pp (p = 0.235) for the high-priors group. Panel B shows that the treatment effect on the share of negotiation attempts is 9.0 pp (p =

[40] Appendix C.5 shows that the findings are similar when the dependent variable is the percent-increase in compensation rather than the share experiencing compensation gains, and it also shows that the treatment had no detectable effects on non-compensation terms.

0.006) for the low-priors group, but only 3.0 pp (p = 0.552) for the high-priors group. And Panel C shows that the treatment effect on the share with compensation gains is 7.3 pp (p = 0.018) for the low-priors group, compared to just 1.6 pp (p = 0.762) for the high-priors group.[41] Moreover, note that for all three outcomes in Figure 4, there is a systematic baseline gap: in the control group, individuals with low priors were less likely to negotiate and to experience compensation gains than those with high priors. By closing the gap in beliefs, the encouragement treatment largely closes the gap in negotiation outcomes. In sum, the large heterogeneity by prior beliefs suggests that the factual content of the encouragement treatment played a central role in driving the observed effects. It also provides reassurance that the treatment effects are genuine rather than spurious.

3.4 Additional Robustness Checks

One potential concern is that, although the treatment increased expected compensation, the additional negotiation attempts it prompted may have come at a cost. To investigate this possibility, Appendix C.5 examines treatment effects on a range of outcomes that could reflect negative repercussions—such as the likelihood of accepting the offer and self-reported measures of job alignment, satisfaction with the negotiation, and perceived backlash. We find no evidence of adverse effects. On the contrary, there is evidence of positive impacts on individuals' perceived alignment with the job offer. Our preferred interpretation is that, during the negotiation process, individuals received—or sought out—additional information about the position, which led them to view it as a better match than they had initially believed.

A separate concern is the potential for selection bias, as some outcomes are measured for a subsample of individuals—specifically, those who responded to the follow-up survey and had at least one offer. If the treatment affected either the response rate or the composition of respondents, this could bias the results. While we observe some differences in response rates across treatment groups (see Appendix C), we find no evidence of composition effects: treatment and control groups remain balanced on pre-treatment characteristics even within the restricted sample of individuals with one offer at follow-up.

[41] The differences between low- and high-priors groups are individually statistically insignificant: p-values of 0.251 in Panel A, 0.326 in Panel B, and 0.328 in Panel C. Nonetheless, when taken jointly, the consistent direction and magnitude of the effects across all three outcomes is reassuring.

We address concerns about selection bias in two main ways. First, we follow the usual approach of reporting Lee bounds alongside each treatment effect. The results are robust. For example, in Panel B of Figure 4, the treatment effect for the "low-priors" group is 9.0 pp, with corresponding Lee bounds ranging from 5.9 to 13.4 pp. The bounds become even tighter—and thus the results more robust—when we apply more sophisticated methods that leverage information on the number of contact attempts. Second, in the style of an event-study analysis, we present falsification tests using pre-treatment negotiation outcomes. Moreover, we estimate these pre-treatment "effects" not only in the full sample, but also in the restricted sub-sample (i.e., who had at least one offer at follow-up). If the post-treatment effects are driven by selection bias, we would expect to observe similar effects in the pre-treatment outcomes. The results, presented in Appendix C.3, go against the hypothesis of selection bias.

Several additional factors suggest that selection bias is unlikely to drive our results. First, selection bias cannot account for the concentration of effects among individuals in the low-priors group. Second, the intention to negotiate is measured in the baseline survey for the full sample, and thus is immune to selection bias. The fact that treatment effects are qualitatively similar between the ex-ante intention to negotiate and the ex-post negotiation behavior further supports the validity of the findings. Lastly, controlling for pre-treatment characteristics—an approach that can mitigate selection bias—yields similar results (Appendix C.4).

4 Effects of the Coaching Treatment

4.1 Willingness to Pay for the Coaching Service

We begin by examining the distribution of stated WTP for the coaching service. The service is designed to equip individuals with the information and skill they need to conduct a successful negotiation. If intensive-margin uncertainty were the primary reason stopping people from negotiating, we would expect a high WTP for the service. While the full price tag may initially seem steep, it can be far outweighed by the potential gains. As a simple back-of-the-envelope calculation, individuals who attempt to negotiate receive, on average, an additional $7,377 in compensation per year. Under conservative assumptions, that translates to a net present value of approximately $15,000.[42] This

[42]This calculation assumes the individual remains at the employer for only three years, that salary gains do not carry over to a new employer, a discount rate of 5%, and a marginal tax rate of 30%.

figure is well above the full price of the coaching service ($1,250-$2,450). Moreover, the fee can be reimbursed if the negotiated raise does not meet the guaranteed minimum. And compensation gains are only part of the potential benefits, which also include reduced anxiety and personalized guidance in choosing between offers.

In contrast, the stated WTP suggests limited demand for the service. Panel C of Figure 3 shows that the median participant has a stated WTP for the negotiation coaching in the range $150–$200—well below the full price of $1,250-$2,450. Although this measure should be incentive-compatible in theory, stated WTP may overstate true demand in practice, as discussed below. This only reinforces our conclusion: if stated WTP is already low, actual demand is likely even lower.

4.2 Average Treatment Effects

Figure 6 presents the results for the coaching treatment arm in Wave 1. Figure 6 is analogous to Figure 4, but for the coaching treatment instead of the encouragement treatment. As before, each panel corresponds to a different outcome. We start with the left half of each panel, which correspond to the results for the full sample. In Panel A of Figure 6, the dependent variable is whether the subject took up the coaching service, regardless of whether they were offered the discount.[43] In the control group, only 1% of subjects took up the coaching service. The discount treatment increased the take-up rate; while the effect is highly statistically significant ($p = 0.001$), it is small in magnitude (1.9 pp). In other words, even at the heavily discounted price of $250, the take-up rate remains very low at 2.9%. The effects on take-up mirror the willingness-to-pay data in indicating that demand for the coaching service is very limited.

In Panel B of Figure 6, the outcome variable is whether the subject, when given the chance, attempted to negotiate. Given the small effect of the offer on take-up, we would expect—if anything—a small effect of the treatment on negotiation attempts. Indeed, the coaching discount increased the share of negotiation attempts by only 0.9 pp ($p = 0.752$). This effect is similar in magnitude to the effect on take-up of the coaching service (1.9 pp, from Panel A).[44] Given the small positive impacts

[43]We use identifying information for each subject—such as full name, email address, and work history—and cross-reference it with the database provided by levels.fyi on all individuals who used their negotiation coaching service. Importantly, this allows us to track take-up of the service regardless of whether the subject received the discount.

[44]The results from these two panels are based on different samples, because Panel A applies to individuals who responded to the follow-up and had at least one offer. In any case, Panel B of Figure C.5 shows that the results from Panel A of Figure 6 are similar when applying the same sample restrictions.

on take-up and negotiation rates, we would also expect—if anything—a small positive impact on the probability of a compensation improvement. Panel C of Figure 6 shows that, indeed, there was no significant increase for this outcome.

To summarize, we establish that among a general population of tech employees, intensive-margin frictions are unlikely to be a major factor preventing individuals from negotiating. This finding stands in stark contrast to the effects of the encouragement treatment, which increases the share of individuals attempting negotiations and experiencing compensation gains.

4.3 Heterogeneity by Willingness to Pay

While access to affordable coaching may not be a key determinant of whether the average individual attempts to negotiate, this result may mask meaningful heterogeneity. There may be a minority of individuals for whom intensive-margin frictions are more significant. To explore this possibility, we can leverage the stated WTP data. In each panel of Figure 6, the right side of the figure presents a breakdown by individuals who reported a WTP below $300—referred to as low-WTP (roughly 25% of the sample)—and those willing to pay above $300, referred to as high-WTP (the remaining 75%).[45] We use $300 as the cutoff because it is right above the discounted price offered to individuals in the treatment group ($250), but the results are similar using alternative thresholds. The law of demand makes a clear prediction: the effects of the coaching treatment on take-up—and consequently on other outcomes—should be stronger for the high-WTP group than for the low-WTP group.

Before diving into the heterogeneity analysis, we first assess the accuracy of the stated WTP data. After all, if stated WTP is pure noise, the heterogeneity analysis would be meaningless. If stated WTP perfectly reflects true preferences, we would expect, in the treatment group (who faced the discounted price of $250), a near-0% take-up among low-WTP subjects,[46] and a 100% take-up among high-WTP subjects. Panel A of Figure 6 shows that the first prediction is close to the data (1.0% take-up versus the predicted 0%), but the second prediction is far from the truth (4.5% take-up versus the predicted 100%). Our preferred interpretation is that stated WTP is subject to "house-money bias": subjects might

[45] More precisely, we split individuals based on their choice in the third scenario, where they had to choose between the coaching service and a $300 gift card.

[46] We say near-0% because some individuals in the low-WTP group may have a WTP between $250 and $300. Nonetheless, the results are similar if we instead focus on individuals with a stated WTP below $200, for whom the predicted take-up is exactly 0%.

report a higher WTP when spending the experimenter's money than they would with their own.[47] Nevertheless, there is still *some* signal in the stated WTP data amid the noise: as expected, take-up is significantly higher in the high-WTP group than in the low-WTP group (4.5% vs. 1.0%, p = 0.010).

We now return to the heterogeneity analysis. As expected, Panel A of Figure 6 shows that the coaching treatment had a stronger effect on take-up in the high-WTP group (4.5 pp, p=0.008) than in the low-WTP group (1.0 pp, p=0.048). However, even within the high-WTP group, the effect on take-up is small, so we would expect correspondingly small (if any) effects on other outcomes such as negotiation attempts or compensation gains. Unfortunately, in part because the high-WTP group represents only 25% of the sample, estimates for this subgroup are imprecise, so we are largely underpowered to detect small effects. Because individuals with strong interest in coaching are relatively rare, studying the effects of coaching on this group would require recruiting a new sample targeted toward this special population. This was the motivation for the second wave of the survey, described below.

4.4 Additional Results: Wave 2

To recruit subjects with the highest demand for the coaching service, we embedded a link to the survey on levels.fyi's website. More precisely, there was a full page dedicated to the coaching service—a screenshot is provided in Figure 2. Right below the button to book the service, we included a second button inviting users to fill out a survey, with the promise that they could qualify for a deep discount.[48]

The subject recruitment started on September 21, 2023, and ended on February 10, 2025.[49] In the second wave, an additional 1,586 subjects completed the baseline survey, 1,057 of whom completed the follow-up survey, among whom 803 received at least one job offer. Relative to the subjects in the first wave, those in the second wave appear broadly similar across many dimensions, including demographic characteristics, years of experience, and the positions and companies they work at. However, there are some notable differences between the two samples (see Appendix C.6). Most importantly, as discussed below, demand for coaching was significantly higher in the second wave

[47] Another source of measurement error in the stated WTP data is subject comprehension—see for example Cason and Plott (2014).

[48] The link was originally only under the mid-level package, and the link under the Senior IC Package was added on October 19, 2023.

[49] We agreed with the partner company to continue recruiting subjects until we reached 300 individuals who claimed the discount, which roughly depleted the remaining budget for this study. The recruitment period was lengthy simply because, on any given day, there were a limited number of people visiting the webpage that contained the link to the study.

than in the first. Another key difference is the negotiation rate. In the control group, 75.5% of subjects in Wave 2 attempted to negotiate, which is substantially higher than the corresponding 53.7% in Wave 1. This difference suggests that while there is still some room for an intervention to increase negotiation rates, the scope is considerably smaller.

The recruitment strategy was successful insofar as subjects in Wave 2 exhibited significantly higher demand for the coaching service. This difference is apparent from their stated WTP. Panel E of Figure C.6 shows that 80% of Wave 2 subjects reported a WTP of $300 or more, compared to just 25% in Wave 1. Most importantly, actual take-up of coaching was much higher in the second wave. In the control group—who faced the full price—12.2% took up the coaching service in Wave 2 (Panel A of Figure 7), compared to just 1.1% in Wave 1 (Panel A of Figure 6). And due to the stronger demand in Wave 2, the discount treatment had a substantial effect on take-up of the service. Panel A of Figure 7 shows that providing the discounted price increased take-up of the coaching service by 24.0 pp ($p < 0.001$), from 12.2% in the control group (facing the full price) to 36.2% in the treatment group (facing the discounted price). Moreover, the effect was stronger among individuals with high-WTP (28.5 pp, $p < 0.001$) than among those with low-WTP (10.4 pp, $p = 0.001$), and the difference between the two groups is statistically significant ($p < 0.001$).

While there was high demand for the coaching service, making the service affordable did not increase negotiation rates, because the marginal individuals who took up the service due to the discount would have negotiated anyways. Panel B of Figure 7 shows the effect on negotiation rates. While a 24.0 pp higher share of individuals took up the coaching service due to the treatment, the corresponding effect on the share attempting to negotiate is positive but small (1.4 pp) and statistically insignificant ($p = 0.648$). Taken together, these estimates imply that 95% ($= \frac{24.0-1.1}{24.0}$)—the vast majority—of the individuals who took up coaching due to the treatment would have attempted to negotiate on their own anyway.[50] Moreover, the results from Figure 7 are precisely estimated—based on the 90% confidence interval, we can rule out an increase in the negotiation rate above 4.7%, meaning that upwards of 80% ($= \frac{24.0-4.7}{24.0}$) of the marginal individuals who took up coaching would have negotiated on their own.

Panel C of Figure 7 shows the effects of the coaching treatment on the share of subjects experiencing compensation gains. The interpretation of these effects is more nuanced, as they may respond

[50]The results from these two panels are based on different samples, because Panel B focuses on individuals who responded to the follow-up and had at least one offer. In any case, Panel C of Figure C.5 shows that the results from Panel A of Figure 7 are similar when applying the same sample restrictions.

to two channels. The coaching treatment may increase the probability of experiencing compensation gains by inducing more individuals to negotiate, but also by helping those who would have negotiated anyways to negotiate more effectively than they would have in the absence of a coach. Panel C shows that the coaching treatment increased the share of subjects experiencing compensation gains, but the effect is small (1.7 pp) and statistically insignificant (p = 0.626). However, this result should be interpreted with caution, as it reflects an intention-to-treat effect, because compliance was far from full—most people who were offered the discount did not end up taking up the service. To approximate a treatment-on-the-treated effect, we can combine the results from Panels A and C. These estimates imply that the marginal person who took up the coaching service because of the discount saw an increase in the probability of a compensation gain of 7.1 pp ($= \frac{1.7}{0.24}$), equivalent to 15% of the baseline probability.

Lastly, while not the central focus of our study, an additional question of interest is whether those who used the coaching service ended up with different outcomes—such as final compensation terms—than if they had not used the service. For space constraints, those additional results are reported in Appendix C.4. A simple comparison of means suggests that, on average, subjects who used coaches experienced a 2.21 pp increase in compensation relative to those who did not. Similarly, those who used the service rated their satisfaction with the negotiation process 0.32 points higher (on a 1 to 5 scale) and were more likely to move to a new employer. However, these findings should be interpreted with caution: although they are precisely estimated, they may suffer from omitted variable bias. In contrast, the corresponding experimental estimates are less precise and therefore sometimes statistically insignificant.

5 Treatment Heterogeneity by Gender

In this section, we examine heterogeneity in treatment effects by gender. This pre-registered analysis is motivated by the fact that negotiations are often cited as a contributing factor to the gender pay gap (e.g., Babcock and Laschever, 2003; Sandberg and Scovell, 2013). Moreover, because the tech sector is a male-dominated industry, women may be particularly susceptible to gender biases. Indeed, we observe several notable baseline differences that may be suggestive of such biases. Compared to men, women are on average 4.9 pp less likely to attempt negotiation and, consequently, 4.4 pp less

likely to experience a compensation improvement. If anything, women have even stronger incentives to negotiate: after controlling for company and position, they receive initial offers that are 5.9% lower than those received by men—see Appendix C.3. Women also differ in other respects, such as believing that negotiations are less common and reporting lower confidence than men—see Appendix B.1.

Any gender differences in the treatment effects should be interpreted with caution. First, men and women may differ in observable and unobservable characteristics besides gender. As a result, the gender differences may be partly or entirely attributed to these other factors. Second, there are power limitations. Since women make up a relatively small share of the sample (19.5% in Wave 1 and 25.6% in Wave 2), the treatment effects for women are imprecisely estimated.

With those caveats in mind, we begin by examining heterogeneity in the effects of the encouragement treatment. Figure 5 mirrors Figure 4, except that the sample is split by the gender of the subjects rather than by their prior beliefs. We find suggestive evidence that while the encouragement treatment increased negotiation attempts—and successful ones—for both men and women, the effects are relatively stronger for women. Panel A of Figure 5 shows that, for the ex-ante intention to negotiate at baseline, there are no significant gender differences in treatment effects. However, Panels B and C reveal that some significant differences emerge later. Panel B indicates that the effect on actual negotiation attempts is positive and statistically significant for men (5.7 pp, $p = 0.062$), and also positive and statistically significant for women, but three times as large (16.8 pp, $p = 0.014$). Panel C shows a similar pattern for compensation gains: a treatment effect of 3.7 pp ($p = 0.201$) for men versus 15.2 pp ($p = 0.024$) for women. While these gender differences in treatment effects are only borderline significant (p-values of 0.141 and 0.110 in Panels B and C, respectively), the consistency in direction and magnitude across both outcomes suggests a meaningful underlying effect.

Next, we examine the heterogeneity in the effects of the coaching treatment. Figure 8 mirror Figure 7, but splits the sample by the gender of the subject rather than by their stated WTP. Panel A of Figure 8 shows that when it comes to the effects of the coaching treatment on take-up, there are no statistically significant differences by gender.[51] In contrast, Panels B and C reveal more pronounced gender differences in the treatment effects on negotiation attempts and compensation gains. Among men, these effects are close to zero and statistically insignificant: -0.2 pp ($p = 0.960$, Panel B) for negotiation attempts and -1.2 pp ($p = 0.754$, Panel C) for compensation gains. In other words, the

[51]The treatment effect on take-up is somewhat larger for men, but the difference is not statistically significant.

men who took up coaching as a result of the treatment were going to negotiate—and experience compensation gains—regardless.

By contrast, most of the women who enrolled in coaching due to the discount would not have attempted to negotiate otherwise—and, as a result, would have been less likely to experience compensation gains. Indeed, this pattern aligns with the perceptions of the coaches themselves, most of whom believe their service is particularly beneficial for women.[52] Panel B of Figure 8 shows that, for women, the coaching treatment had a large positive effect on the probability of negotiating (9.9 pp), though the estimate is borderline insignificant ($p = 0.174$). Consistently, Panel C shows a positive effect on the likelihood of experiencing compensation gains (12.5 pp), although this effect is again borderline insignificant ($p = 0.125$). While each estimate is individually statistically insignificant, the fact that they are both consistent in direction and magnitude is reassuring. Moreover, the effects of the coaching treatment on women (Panels B and C of Figure 8) closely mirror those of the encouragement treatment (Panels B and C of Figure 5). In both cases, the interventions increased the likelihood that women would negotiate, and these additional negotiations were largely successful. Thus, even if each individual estimate is not statistically significant, the consistency of the patterns across treatments provides stronger evidence of meaningful gender differences.

6 The Model

6.1 Outline of the Model

Our field experiment on salary negotiations reveals several key insights. To capture empirically grounded information frictions in negotiation, we develop a theoretical framework with three essential building blocks:

1. Employees have bargaining power: Employees are not passive price-takers in the labor market, and can choose to make counteroffers to employers.

2. Information asymmetry about offer negotiability: Employees do not necessarily know if an offer is "negotiable," which implies that employers cannot credibly claim an offer is "best

[52] Among the 12 coaches who responded to our survey, 8 believed that the coaching service would be more beneficial for women, while only 1 believed it would benefit men more.

and final." An employee who attempts to negotiate a non-negotiable offer experiences negative consequences. This creates a risk to negotiating that employees must consider, distinct from prior incomplete information models that focus on intensive-margin frictions rather than extensive-margin uncertainty.

3. Multiple offers: Employees potentially have multiple job offers, making their fallback option endogenous to negotiation decisions with other employers. In other words, employees have an "inside option" which varies based on their negotiation decisions with other employers.

These features have important implications for efficiency in the labor market, because asymmetric information can lead to inefficient outcomes. Similarly to Akerlof (1970)'s lemons model, information frictions can prevent matches between an employee and an employer with a high value of labor from forming—employees may fail to negotiate when they should have, ending up at a lower value employer. On the other hand, employees may push for higher compensation when they shouldn't have and lose a match entirely. An important distinction from classic models is that with multiple offers, employees face a portfolio optimization problem (as in Chade and Smith, 2006) because they each seek only one job. This implies that the decision to negotiate with one employer is interdependent both to the other offers the employee receives, but also on her choices to negotiate these other offers.

We first consider the negotiation decision of an employee who receives a portfolio of job offers and can elect to negotiate with any of the offering employers by making a counteroffer before selecting a job. We characterize the employee's optimal bargaining strategy as a function of her portfolio of offers and her beliefs.

Our model, designed to capture the key information frictions from our experiment, yields striking predictions for negotiation patterns in observational data that distinguish it from other bargaining models applied to the labor market. For example, our model predicts employees demand a premium above and beyond their alternative options, and that the act of negotiation generates dispersion in final offers. These predictions stand in stark contrast to models based on Bertrand competition that lead to offer equalization through bargaining, and those based on Nash (or alternating) offer bargaining which predict no employee counteroffers on equilibrium path. We describe broad support for our model in the observational data.

Finally, we use our model to predict the effects of policies that promote (or discourage) negotiation.

Because these policies affect the negotiation of employees market-wide, employers may change their wage offers in equilibrium. Theoretical analysis of equilibrium effects is valuable because our field experiment considers the effects of changing the negotiation behavior of a small fraction of the workforce, and thus, does not include the indirect effects of employer responses. We show that inducing employees to "lean in" and negotiate with high probability has positive effects on efficiency, employee surplus, and pay equity between employees with similar job offers. Moreover, we show that a lean-in equilibrium can plausibly be induced by a "scaled up" version of our experimental encouragement treatment, i.e., a messaging campaign that tells all employees in the labor market that it is safe to negotiate.

6.2 Setup: Single Employee

There is a single employee with an outside option denoted $\varnothing$ with value (or "wage") θ normalized to zero. There is a set of employers $N := \{1,2,...,N\}$, that each make the employee a job offer.[53] For ease of exposition, we refer to both the set of employers and its cardinality by N.

As in Abowd et al. (1999), the employee generates a commonly known match surplus $v_i \in \mathbb{R}$ if matched with employer $i \in N$,[54] where we assume that $v_i \neq v_j$ for all distinct employers $i,j \in N$. Each employer i offers a wage $w_i \leqslant v_i$.

Upon receiving her portfolio of offers $\{w_i\}_{i \in N}$ the employee engages in simultaneous negotiations by making counteroffers. Formally, the employee simultaneously selects an additional wage demand $o_i \geqslant 0$ for each employer $i \in N$; if $o_i > 0$ then we say that the employee makes a counteroffer to employer i. Making a counteroffer to employer i potentially exposes the employee to negative consequences for two reasons. First, i's offer may be *non-negotiable* in which case making a counteroffer results in i's job offer being withdrawn. Second, even if i's offer is negotiable, i will reject a counteroffer and withdraw the offer if the employee demands more than the match value, $o_i > v_i - w_i$. We say that an employee *negotiates with employer* $i \in N$ if either $v_i = w_i$ (the employer initially capitulates and offers the employee the entire match value, thus obviating the need for the employee to make a counteroffer) or $o_i > 0$ (the employee makes a counteroffer).

[53] We can interpret a currently employed worker as one with an additional employer in her portfolio of offers. Our model straightforwardly accommodates cases in which the worker can or cannot initiate negotiations with her current employer.

[54] We do not take a stance on the genesis of each match value v_i, and treat it as exogenous. It represents the entire surplus generated from the worker's employment at employer i if the employer has no prospect of filling the position with another worker, or if employer i has an outside option of instead hiring a part-time worker at a known rate, v_i would represent the surplus over and above this outside option.

The employee has *extensive-margin* uncertainty, captured by vector $\hat{\rho} = (\hat{\rho}_1, ..., \hat{\rho}_N)$, where $\hat{\rho}_i$ specifies the employee's belief that employer i's offer is negotiable. For each employer $i \in N$ let $\tilde{w}_i$ be the final surplus of matching with employer i:

$$\tilde{w}_i = \begin{cases} w_i & \text{if } o_i = 0, \\ -\infty & \text{if } o_i > v_i - w_i, \text{ and} \\ \begin{cases} w_i + o_i & \text{with probability } \hat{\rho}_i, \\ -\infty & \text{with probability } (1-\hat{\rho}_i) \end{cases} & \text{if } o_i \in (0, v_i - w_i] \end{cases}$$

Recalling that the wage offered by the outside option $\varnothing$ satisfies $\tilde{w}_\varnothing = 0$, the worker is hired by an employer $i^* \in \underset{i \in N \cup \{\varnothing\}}{\text{argmax}}\, \tilde{w}_i$ and receives a final payoff equal to her wage $\tilde{w}_{i^*}$.

This assignment procedure has the following interpretation. The employee accepts a job at (one of) the employer paying the highest wage. The wage an employer pays is equal to the initial wage offer ($\tilde{w}_i = w_i$) if $o_i = 0$. If $o_i > v_i - w_i$, then this offer drops in value below that of the employee's outside option ($\tilde{w}_i = -\infty$) because the offer is rescinded.[55] Otherwise, the employer accepts the terms of the employee's counteroffer ($\tilde{w}_i = w_i + o_i$) if the original offer was negotiable and rescinds the offer ($\tilde{w}_i = -\infty$) if the original offer was non-negotiable.

Remark 1. *We have described a game in which all negotiations occur simultaneously, and an employer's acceptance of an employee's counteroffer does not oblige the employee to matriculate at that employer. More realistically, job offers are negotiated sequentially, and negotiating plausibly commits the employee to accepting that job, i.e., the employee may need to declare, "if you accept my counteroffer, I'll sign on the dotted line right now," in order for the employer to take the counteroffer seriously.*

The solution to our simultaneous-negotiation game is equivalent to that of a sequential-negotiation game in which the employee selects an order in which to negotiate with the employers, and she is matched to the first employer that either accepts her counteroffer, or to whom she makes no counteroffer (i.e., she accepts the employer's initial offer).

Remark 2. *We have assumed that a failed negotiation results in offer rescission. An alternative assumption is that that making a counteroffer exposes an employee to a distribution of negative*

[55] Therefore, the normalization that $\tilde{w}_i = -\infty$ if the offer is rejected is merely a mathematically convenient way to say the employee will never be assigned to an employer that rejects her counteroffer.

consequences, reflecting frictions in the negotiation process. Specifically, a failed negotiation may result in the initial offer still being available to the employee, but some of the match surplus may be destroyed due to a deterioration of the relationship. Additionally, a negotiation (whether successful or not) may result in costly delays in onboarding. Our results are qualitatively similar if we embed these possibilities as follows: rejected counteroffers result in the offer either being withdrawn or having the offer remain constant and the match value lowered to w_i, each with positive probability; "accepted" offers are only available to the employee probabilistically, reflecting the time cost of negotiating.

6.3 Optimal portfolio bargaining

We seek a solution which maximizes the employee's expected utility:

1. For a given portfolio of wage offers $\{w_i, v_i, \hat{\rho}_i\}_{i \in N}$, optimal counteroffers $\{o_i^*\}_{i \in N}$ maximize the employee's expected (given $\hat{\rho}$) payoff, and

2. If the employee has a zero probability of matching with an employer $i \in N$ according to criterion 1, then $o_i^* = 0$.[56]

An initial observation is that if an employee negotiates with an employer $i \in N$ then the optimal counteroffer equals $v_i - w_i$; demanding the entire match surplus maximizes the wage available at employer i conditional on the offer being negotiable, because any lower counteroffer is accepted by the employer, while any higher counteroffer ensures rejection.

Remark 3. *The optimal counteroffer for each employer* i *satisfies* $o_i^* \in \{0, v_i - w_i\}$.

The optimal counteroffer portfolio trades off risk—captured by $\hat{\rho}_i$ values—and reward—captured by $v_i - w_i$ values. Intuitively, the larger is $\hat{\rho}_i$ and the smaller is $v_i - w_i$, the less likely the employee is to make a counteroffer to employer i (i.e., she is more likely to set $o_i^* = 0$). For extreme values, this decision is simple: if $\hat{\rho}_i = 0$ then the employee is not willing to make a counteroffer to employer i regardless of the value $v_i - w_i$, because the employee perceives no chance of success. On the other

[56] There is a multiplicity of optimal counteroffers if the employee has zero probability of matching with employer i. We restrict the employee to break her indifference by not making a counteroffer to such an employer. In a richer model in which the employee faced a positive cost for negotiating with an employer, optimality would require $o_i = 0$ for any such employer i.

hand, if $\hat{\rho}_i = 1$ then the employee is willing to negotiate with employer i regardless of the value $v_i - w_i$, because the employee perceives no risk of doing so.

For values of $\hat{\rho}_i \in (0,1)$ an additional complication arises when constructing the optimal counteroffer to employer i: the risk versus reward calculation is affected by the presence of other offers, and the endogenous negotiation decisions the employee makes for these offers. Intuitively, even given a relatively low value of both $\hat{\rho}_i$ and $v_i - w_i$, the employee will be willing to negotiate with employer i if she has a solid backup offer. For example, consider the case in which the employee has another offer from an employer $j \neq i$ with $\hat{\rho}_j = 0$ and $w_j = w_i - \epsilon$, for some small $\epsilon > 0$. In this case, the employee will optimally not negotiate with employer j because $\hat{\rho}_j = 0$, so the presence of j's initial offer serves as a risk-free alternative. The employee will optimally "roll the dice" and negotiate with employer i since her effective outside option w_j is only slightly worse than what she gets from not negotiating with i. On the other hand, for any $\hat{\rho}_i < 1$, if w_i is large enough and $v_i - w_i$ is small enough, then the employee will not be willing to risk the potential of losing the job offer by negotiating if she has no or low offers from other employers. The remaining difficulty in characterizing the optimal counteroffers is that the reward from negotiating is endogenous to other negotiation decisions. That is, if the employee risks it by negotiating with an employer i, she may optimally decide to not negotiate with an employer j and instead keep it as a back up option.

These complexities are similar to those present in Doval (2018), in which an agent searches by opening "boxes" at a cost to reveal their values, before making a final selection. Unlike in Weitzman (1979)'s seminal model, the agent can select a box without opening it, and the decision to open a box depends on the set of boxes available to the agent. In our setting, "opening a box" is analogous to making a counteroffer, where the cost is the probabilistic withdrawal of the offer, and choosing not to "open a box" is analogous to settling for the employer's initial wage offer. In both settings, optimal behavior does not follow a (Gittins') index policy, specifically because of the interdependency of optimal choices on the entirety of the portfolio. We proceed to construct the optimal counteroffers given the entire portfolio of offers in an inductive manner.

The following claim states that the employee may optimally negotiate with any subset of employers in her choice set. In other words, an analyst who observes an employee's collection of initial offers can rationalize negotiations with any subset of employers in her portfolio by some collection of match values and beliefs.

Claim 1. *Let* N *be a collection of employers, with initial offers* $\{w_i\}_{i\in N}$. *Let* N' *be an arbitrary, non-empty subset of* N. *Then there exists a collection of match values and beliefs* $\{v_i,\hat{\rho}_i\}_{i\in N}$ *such that the employee optimally negotiates with all employers in the set* N' *and no employer in the set* $N\setminus N'$.

Reflecting Claim 1, Example 1 in Appendix A.1 demonstrates conditions under which it is optimal for the employee to negotiate with each subset of the two employers making her job offers.

An optimal portfolio of counteroffers always exists despite these complex interdependencies. Moreover, the following proposition finds that for (almost) any collection of parameter values, there is a unique optimal vector of counteroffers. Its proof also implies Claim 1.

Proposition 1. *For any given portfolio of offers, there generically (in the space of match-value vectors* v*) exists a unique optimal vector of counteroffers.*

We provide an algorithm in the appendix which finds the optimal portfolio of counteroffers, using a double induction argument to build the optimal portfolio from the "bottom up." One induction loop constructs the optimal counteroffer to a particular employer i given the employee's expected payoff from employers m with $v_m < v_i$ *given the already-calculated counteroffers* o_m. Our construction demonstrates that each employer m with $v_m < v_i$ serves as an "inside option" that the employee considers when negotiating with employer i, where the value of the inside option depends on the values of v_m and w_m (the reward), and on $\hat{\rho}_m$ (the risk). Under the inductive hypothesis that the employee wishes to negotiate with all employers from the "bottom up" as described above, then we have found the optimal portfolio of counteroffers. If not, then we pause the first induction loop at the first employer i' where our inductive hypothesis breaks down, meaning that the employee is not willing to make a counteroffer to employer i' given her inside option from lower-value employers, and would rather keep employer i' as a safety option given the counteroffers constructed to this point.

In case of a breakdown of our first induction loop with an employer i, our second induction loop recalculates optimal counteroffers for employers m with $v_m < v_i$ under the assumption that the employee does not negotiate with i. In general, the induction loops can "interrupt" each other multiple times, because, even if the employee was not willing to make a counteroffer to some employer i given the constructed counteroffers to lower-value employers, if she later elects not to negotiate with some employer j with $w_j \in (w_i,v_i)$, she will want to revise her decision not to negotiate with i since she now has a higher-wage safety option. However, by negotiating with employer i, the employee lowers

her "inside option" when she considers negotiating with employer j. Nevertheless, we show that an important monotonicity condition obtains: the employee will still not negotiate with j if her inside option from lower employers is decreased. This importantly means that any time the second induction loop starts, it has advanced by at least one employer. Therefore, the algorithm must terminate, and does so at the optimal portfolio of counteroffers.

Based on Proposition 1, we henceforth assume the existence of a unique optimal vector of counteroffers, and study its properties. First, Remark 3 implies the employee will demand a different wage from each employer she makes a counteroffer to, and the following claim additionally implies that an employee will initiate negotiations with (all but one, or all) employers whose initial offers are equal if she perceives a strictly positive probability that the offer is negotiable. In other words, the employee is not indifferent between any two job offers after negotiating.

Claim 2. *After negotiating, the employee has strict preferences over all non-rescinded job offers in her portfolio:* $w_i + o_i^* \neq w_j + o_j^*$ *for all* $i \neq j$ *such that* $\hat{\rho}_i \neq 0$ *and* $\hat{\rho}_j \neq 0$.

Second, optimal counteroffers follow a "top down" structure. Supposing, without loss of generality, that $w_1 + o_1^* > w_2 + o_2^* > ... > w_N + o_N^*$, the employee will be employed at employer 1 if and only if her counteroffer to employer 1 is not rejected, she will be employed at employer 2 if and only if her counteroffer to employer 1 is rejected and her counteroffer to employer 2 is not rejected, and so on. Because of this, we show that the employee will make counteroffers to the employers with the highest values, as characterized by the following result.

Proposition 2 ("Top down" monotonicity and Leverage)**.** *Fix a collection of initial offers, negotiation beliefs, and match values* $\{w_i, \hat{\rho}_i, v_i\}_{i \in N}$ *such that* $\hat{\rho}_i \in (0,1)$ *for all* $i \in N$.

1. *Let* j *be the employer with the highest initial offer to which the employee optimally does not make a counteroffer, i.e.,* $o_j^* = 0$ *and* $o_i^* > 0$ *for all* $i \neq j$ *such that* $w_i \geqslant w_j$. *Then the employee optimally negotiates with every employer* ℓ *such that* $v_\ell > v_j$.

2. *Suppose* $w_i < v_i$, *and let* θ_i *represent the expected surplus the employee receives in her optimal portfolio if she were rejected from all employers* j *such that* $w_j + o_j^* \geqslant w_i + o_i^*$. *Then she optimally makes a counteroffer to employer* i *if and only if* $\theta_i \geqslant \frac{w_i - \hat{\rho}_i v_i}{1 - \hat{\rho}_i}$.

Point 1 describes the "top down" structure of optimal counteroffers we have previously discussed. Point 2 relates this top down structure to leverage. It finds that the "threshold" employer j in Point 1 is the employer with the highest initial offer such that a local optimization constraint fails: the employee finds the risk of negotiating too high compared to the inside option (represented by θ_i) given the remaining portfolio optimal counteroffers. The inside option θ_i can be thought of as a notion of leverage; the higher θ_i, the less the employee has to lose from negotiating with employer i. Point 2 therefore states that the higher the inside option, the more leverage the employee has in a negotiation, which increases her desire to negotiate.

6.4 Empirical Validation of the Model

Before exploring the model's equilibrium implications, we compare its predictions to the data and to the predictions of other bargaining models that have been applied to the labor market.

Our key experimental finding is that the encouragement treatment increases the share of employees who attempt to negotiate and secure compensation gains. Our model can rationalize this finding.[57] In sharp contrast, complete-information bargaining models (Mortensen and Pissarides, 1994; Gentile Passaro et al., 2024) cannot account for this key finding. These models assume all relevant information in the labor market is common knowledge. Hence, they predict that information about negotiation rates in the broader market would not affect bargaining between an employee and an employer. And while there are some bargaining models with incomplete information (e.g., Myerson and Satterthwaite, 1983; Chatterjee and Samuelson, 1983; Fudenberg et al., 1985; Fuchs and Skrzypacz, 2010), they do not have uncertainty at the extensive-margin and thus cannot rationalize the experimental finding either.[58]

We can also contrast our model (and the alternative ones) to some basic non-experimental facts. One basic fact is that employees often exercise their bargaining power by making counteroffers. This is a central prediction of our model, according to which the employee weights risk versus reward

[57] More formally, as we discuss in Claim 3 in Appendix A.2, employees with large initial misperceptions are predicted by our model to experience the largest gains in propensity to negotiate and increase in expected wages.

[58] The bargaining literature with one-sided incomplete information separately studies the "gap" case (i.e., the uninformed party knows there are gains from trade available) from the "no gap" case (i.e., the informed party believes with positive probability that there are no gains from trade). See Ausubel et al. (2002). Both of these differ from our approach. The "gap" case does not embed extensive-margin uncertainty, in that the uninformed party knows she can demand more surplus. The "no gap" case has both intensive- and extensive-margin uncertainty; the uninformed party does not know how much she can demand, or if she can even demand any surplus.

when deciding to negotiate each offer, and as a result, the employee may optimally negotiate with any subset of employers making her offers (Claim 1). This fact goes against a family of bargaining models derived from seminal models by Nash (1950) and Rubinstein (1982), in which employees individually negotiate with employers (e.g.s, Mortensen and Pissarides, 1994; Gentile Passaro et al., 2024). Contrary to the data, these models instead predict that we should observe "passive" employees; an employee who takes a new job is never predicted to demand more than that employer's initial offer. Indeed, recent empirical papers studying bargaining also use the observance of attempts to negotiate initial offers as evidence against this family of models (Backus et al., 2020; Caldwell et al., 2024). A second family of models assumes that employees do not play an active role in wage setting, and that employers instead engage in Bertrand-like competition over workers (e.g., Postel-Vinay and Robin, 2002; Cahuc et al., 2006). In principle, such models could rationalize the observation of employees making counteroffers if negotiation simply involves forwarding outside offers to competing employers, who then match the terms. To distinguish our model from this class of models, we examine predictions about which offers are negotiated and which are revised.

Our data show that employees are more likely to negotiate their best offers than their worst, and that negotiations increase the dispersion of final compensation terms—for more details, see Appendix C.7. This evidence aligns more closely with our model, which predicts that employees begin by negotiating their top offers, treating others as fallback options—thereby increasing compensation dispersion. This stands in sharp contrast to Bertrand-style models, which predict that employees negotiate their lowest offers to match their best, resulting in reduced dispersion in final compensation. Moreover, the evidence suggests that firms rarely matched the compensation terms of competing offers.[59]

6.5 Setup: Equilibrium

Thus far, our theoretical analysis has considered the perspective of a job seeker in the economy. In this section, we embed this single-employee decision problem into a setting where employers make strategic decisions on the terms of initial job offers and whether or not to allow negotiations. This allows us to consider equilibrium effects of policies and messaging campaigns that promote or discourage negotiations.

[59] Offer matching may be the predominant form of bargaining in other marketsâmost notably, universities are known for adjusting pay to match competing offers in academia.

Naturally, not all employees face the same fundamentals: employees differ in their observable characteristics (e.g., gender), in their unobservable characteristics (e.g., outside option), and in their success at search (e.g., number of job offers). We first introduce the setup and timing of our model, and then describe the interactions that our model allows.

An employee in our model is associated with a type composed of a group identity, either A or B, an outside option $\theta \in \Theta$ where Θ is a compact subset of $\mathbb{R}$, a subset of employers $N \subset \mathcal{N}$ she receives offers from (we assume $|N| \geqslant 1$ to rule out trivial cases), and a belief negotiation probability vector $\hat{\rho} \in [0,1]^{\mathcal{N}}$.[60] Let $\underline{\theta}$ and $\bar{\theta}$ represent the minimum and maximum outside options, respectively, in Θ. We represent a generic employee type as $r \in \{A,B\} \times \Theta \times 2^{\mathcal{N}} \times [0,1]^{\mathcal{N}}$. All employee types have a common match value ν_i for each employer $i \in \mathcal{N}$. That is, we are considering a set of employees that are equally productive across employers, but are heterogeneous across other dimensions.

First, each employer $i \in \mathcal{N}$ commits to initial wage offers, and to a negotiation policy. Because we are considering a scenario in which employers observe an employee's group identity, we allow the initial wage offer of each employer to depend on the employee's group identity, which we denote by w_i^g, $i \in \mathcal{N}$ and $g \in \{A,B\}$, thus allowing for the possibility of "wage offer discrimination." Simultaneously, each employer $i \in \mathcal{N}$ selects a probability ρ_i of entertaining counteroffers. We do not allow this probability to differ by group identity; mechanically, this means that we are not allowing for "negotiation discrimination," however, as will become clear shortly, such a constraint does not affect our upcoming analysis.

Second, Nature selects a single employee to enter the labor market, given a commonly known distribution F over employee types. We place little structure on this distribution, which allows us to explore correlations between different elements of the employee's type, e.g., it could be that $B-$group employees typically have low outside options. Similarly, distribution F captures the network of employer competition; it may be the case that with high probability either both employers i and j appear in an employee's portfolio, or neither appears. Our only requirement is that distribution F has full support over all $(g,\theta,N) \in \{A,B\} \times \Theta \times 2^{\mathcal{N}} \setminus \{\emptyset\}$—there is positive probability of drawing an employee of either group, with an outside option of (approximately) any value, who receives offers from any non-empty subset of employers.

[60] Note that we consider "behavioral" employees whose beliefs are exogenously given. This assumption is made to comport with our empirical setting in which employees have misspecified beliefs about offer negotiability. We will separately consider equilibria in which employees have accurate versus misspecified beliefs.

Third, the employee selected by Nature engages in the single-employee search problem described above.: given the selected group identity g, employer subset N, outside option θ, and belief vector $\hat{\rho}$, the employee selects the optimal counteroffers for portfolio $\{w_i^g - \theta, \hat{\rho}_i, v_i - \theta\}_{i \in N}$.

Letting w represent the employee's final pay (defined as $w = \theta$ if the employee is assigned to her outside option), she receives a payoff of $w - \theta$, and the payoff each employer i receives is zero if it does not employ the employee, and $v_i - w$ if it does.

Collecting the primitives described above, we parameterize a *labor market* by $(\mathcal{N}, \Theta, F, \{v_i\}_{i \in \mathcal{N}})$. We continue to assume for ease of exposition that $v_i \neq v_j$ for all distinct $i, j \in \mathcal{N}$. We now introduce our solution concepts.

Definition 1 (Equilibrium)**.** *Consider a labor market* $(\mathcal{N}, \Theta, F, \{v_i\}_{i \in \mathcal{N}})$. *The following three conditions will be used to define our solution concepts:*

> **Employee optimization** *Each employee type* r *of group* g *in the support of* F *solves the optimal portfolio problem* $\{w_i^g - \theta^r, \hat{\rho}_i^r, v_i - \theta^r\}_{i \in N^r}$,
>
> **Employer optimization** *Each employer* $i \in \mathcal{N}$ *sets* $w_i^A \in [-\bar{\theta}, v_i]$, $w_i^B \in [-\bar{\theta}, v_i]$, *and* $\rho_i \in [0, 1]$ *to maximize its expected payoff given* $(\mathcal{N}, \Theta, F, \{v_i\}_{i \in \mathcal{N}})$ *and* ***Employee optimization****, and*
>
> **Rational expectations** *For each employee type* r *in the support of* F, $\hat{\rho}_i^r = \rho_i$.

We say that the labor market is in partial equilibrium *if it satisfies* ***Employee optimization*** *and* ***Employer optimization****. We say that the labor market is in* general equilibrium *if it satisfies* ***Employee optimization****,* ***Employer optimization****, and* ***Rational expectations****.*

The first two conditions laid out in Definition 1 are standard optimization conditions: all employers and employee types maximize their expected utility given their beliefs. Partial equilibrium captures our empirical finding that some employees' beliefs regarding negotiation risk are misspecified: we allow $\hat{\rho}_i^r \neq \rho_i$. By contrast, rational expectations additionally requires belief consistency. Mechanically, employee beliefs are determined through the support of F—a labor market in general equilibrium involves a distribution F that never selects "incorrect" beliefs. Therefore, in general equilibrium (but not in partial equilibrium) each employer i's choice on offer negotiability potentially deters negotiations. Henceforth, we use the unqualified term "equilibrium" to mean either a partial or general equilibrium.

However, even a general equilibrium may not feature negotiation *rates* that match employee beliefs. It may be the case that a particular employer i selects a high value ρ_i, but it may be that only a small fraction of employee types actually elect to negotiate with this employer. A regulator who does not know all details of the labor market (e.g., the outside option of each individual employee) but nevertheless seeks to change the equilibrium through a campaign to affect employee beliefs (e.g., by encouraging all employees to negotiate more) may experience difficulties in doing so given a mismatch between employee beliefs and actual negotiation rates. That is, if an employee learns that very few of her peers actually negotiate with employer i, (the regulator may be concerned that) she will be swayed to believe that ρ_i is smaller than what the regulator messages, causing her not to negotiate with i. The following definition states what it means for a belief vector to be immune from such concerns.

Definition 2. *For any* $(\mathcal{N},\Theta)$ *we call a collection of beliefs* $\hat{\rho} \in [0,1]^{\mathcal{N}}$ robustly self-enforcing *if for all labor markets* $(\mathcal{N},\Theta,F,\{v_i\}_{i\in\mathcal{N}})$ *such that* $\hat{\rho}$ *is the only belief vector in the support of* F*, there exists a general equilibrium in which the probability of negotiation between the employee and each employer* $i \in \mathcal{N}$ *is* $\hat{\rho}_i$.

The following result states the existence of partial and general equilibria, and general equilibria with robustly self-enforcing beliefs. Each of these solution concepts is more restrictive than the last: a partial equilibrium exists for any collection of employee beliefs, a general equilibrium exists if all employees have the same "correct" beliefs, and robustly self-enforcing beliefs require employees to believe each employer i makes a deterministic decision on whether or not to accept counteroffers.

Proposition 3.

1. *Any labor market* $(\mathcal{N},\Theta,F,\{v_i\}_{i\in\mathcal{N}})$ *admits a partial equilibrium,*

2. *A labor market* $(\mathcal{N},\Theta,F,\{v_i\}_{i\in\mathcal{N}})$ *admits a general equilibrium if and only if there is a single belief vector* $\hat{\rho}$ *in the support of* F*,*

3. *Any* $(\mathcal{N},\Theta)$ *admits a robustly self-enforcing belief vector* $\hat{\rho} \in [0,1]^{\mathcal{N}}$ *if and only if* $\hat{\rho}_i \in \{0,1\}$ *for all* $i \in \mathcal{N}$.

Intuitively, a partial equilibrium always exists for any belief vector $\hat{\rho}$ chosen by F because employee beliefs and employer decisions to allow counteroffers are divorced. A general equilibrium

always exists given that the "correct" belief vector $\hat{\rho} = \rho$ is always chosen, because employers' decisions to allow counteroffers only affects their payoffs through deterring employees from making a counteroffer: a successful negotiation leaves the employer with zero surplus, but so does an unsuccessful one. Robust self-enforcement requires employee certainty in the negotiation outcome, so that even employees with extreme risk versus reward considerations will not be deterred from following the regulator's recommendation. To demonstrate why certainty in offer negotiability implies robust self-enforcement of beliefs, consider $\hat{\rho}_i^r = \rho_i = 1$ for all i and all employee types r. Then there is a "lean-in" equilibrium in which all employers negotiate by capitulating all of the surplus to employees, $w_i = v_i$ for all i.[61] Then, regardless of the the specifics of the labor market, the equilibrium negotiation rate at each employer, 1, matches the probability of offer negotiability.

6.6 Comparison of Equilibria

We study how different equilibria result in different labor-market outcomes. Consider a non-empty family H of labor markets $\{(\mathcal{N},\Theta,F^h,\{v_i\}_{i\in\mathcal{N}})\}_{h\in H}$ where for each $h \in H$, P^h is the set of belief vectors in the support of F^h, and the marginal distributions over all other components are constant across all $h \in H$. In words, we consider a family of labor markets which differ only in the belief vector of employees. Recalling Proposition 3, this formalization allows us to consider a family of equilibria that vary only in their belief vectors; let E^h represent the set of equilibria of the labor market indexed by h. In what follows, all probability statements are taken with respect to the distribution of primitives F^h, the equilibrium $e^h \in E^h$, and the negotiation probability vector ρ selected by employers in equilibrium e^h.

We define the following notation within this family of labor markets and their respective sets of equilibria. For any labor market $h \in H$ and any employee type r in the support of F^h with outside option θ^r and offers from employers $N^r \subset \mathcal{N}$, let $V^{N^r} := \max\left\{\{v_i : i \in N^r\} \cup \{\theta^r\}\right\}$ represent the highest achievable match value given offers from the set of employers N^r, including her outside option. For a given $N \subset \mathcal{N}$, let $R^{N,g}$ represent all employee types in the support of F^h that encounter set $N^r = N$ of employers, are of group $g \in \{A,B\}$, and whose efficient match in their portfolio is not their outside option, i.e., there exists $i \in N$ such that $V^N = v_i$.[62] Let $w^{e^h,r}$ and $v^{e^h,r}$ represent

[61] As a result, the lean-in equilibrium may have additional efficiency gains by reducing other negotiation costs associated with making counteroffers that are not included in the model, such as the psychological costs (e.g., anxiety) and recruiter's time.

[62] The restriction to employees whose efficient match is not their outside option ensures that we consider employees

employee r's expected wage and the expected match value r generates, respectively.

- The *efficiency* of an equilibrium e^h is equal to the expected fraction of the maximum match value generated: $\mathbb{E}(v^{e^h,r}/V^{N^r})$.
- The *employee surplus* in an equilibrium e^h is equal to the expected fraction of the maximum match value captured by the employee: $\mathbb{E}(w^{e^h,r}/V^{N^r})$.
- The *wage gap between groups* in an equilibrium e^h is equal to the expected absolute difference in wages between employees from different groups with job offers from the same employers:

$$\sum_{N\subset\mathcal{N}} Pr(N)\cdot\left|\mathbb{E}\left(w^{e^h,r}:r\in R^{N,A}\right)-\mathbb{E}\left(w^{e^h,r}:r\in R^{N,B}\right)\right|,$$

  where $Pr(N)$ represents the probability that $N\subset\mathcal{N}$ is the subset of employers the selected employee type encounters which is common across all F^h, $h\in H$.[63]

Importantly, the above notions hold fixed the set of employers making the employee offers, which means our notion of wage gap between groups compares "similar" employees (in terms of job offers) from different groups. In other words, our measure is potentially subject to Simpson's Paradox: one equilibrium may reduce the wage gap between groups compared to another equilibrium, but may not reduce the "uncontrolled" wage gap in the market overall if $A-$group and $B-$group employees encounter some subset of employers N with different probabilities.

The following proposition presents an evaluation of lean-in general equilibria in which $\hat{\rho}_i=1$ for all $i\in\mathcal{N}$. All such general equilibrium are payoff equivalent, and compared to any other equilibrium outcome, a lean-in general equilibrium yields higher efficiency, higher employee surplus, and smaller wage gap between groups. Indeed, any lean-in general equilibrium maximizes the efficiency and employee surplus, and minimizes the wage gap between groups, within the set of outcomes that yield no employer a strictly negative profit. Also recall from Proposition 3 that the belief vector associated with lean-in general equilibria is robustly self enforcing, i.e., it can plausibly be induced by public policy.

---

who are active in the labor market; an employee r for whom $\theta^r>v_i$ for all employers $i\in N^r$ will never match with a "real" employer, regardless of the selected equilibrium.

[63] We note that our upcoming theoretical results hold under alternative functional forms for measuring the wage gap between groups. For example, there are no changes if we instead defined the wage gap between groups as equaling the maximum expected wage gap for employees of different groups across all subsets $N\in\mathcal{N}$.

Proposition 4 (Efficacy of lean-in general Equilibria)**.** *Let* $(\mathcal{N},\Theta,F,\{\nu_i\}_{i\in\mathcal{N}})$ *be a labor market such that such that* $\hat{\rho}_i = 1$ *for all* $i \in \mathcal{N}$ *is the unique belief vector in the support of* F*. Then in any general equilibrium of this labor market: the efficiency and employee surplus is 1, while the wage gap between groups is 0.*

Proposition 4 is not a knife edge result—any general equilibrium with a belief vector $\hat{\rho} \approx \vec{1}$ achieves nearly the same properties as lean-in general equilibria featuring $\hat{\rho} = \vec{1}$.

Proposition 5. *Consider an infinite sequence of labor markets* $\{(\mathcal{N},\Theta,F^h,\{\nu_i\}_{i\in\mathcal{N}})\}_{h=1,2,...}$ *where for each* h*, there is a single belief vector* $\hat{\rho}^h$ *in the support of* F^h *and* $\hat{\rho}^h \to \vec{1}$ *in* h*. For any* $\epsilon > 0$*, there exists some* h^* *such that for all* $h > h^*$ *and in any general equilibrium* $e^h \in E^h$*, the efficiency and employee surplus is no less than* $1-\epsilon$*, while the wage gap between groups is no more than* ϵ*.*

Our results in Proposition 4 and Proposition 5 apply *within* a labor market, while another policy-relevant consideration is the impact of switching from one equilibrium to another *across* labor markets. Our model offers some guidance on the "elasticity" of outcomes across labor markets: we expect outcomes to meaningfully vary across general equilibria in markets that are imperfectly competitive, not those that are perfectly competitive. First, consider an imperfectly competitive market with a known, common outside option for employees (i.e., Θ is a singleton set), and a single monopsonist i with a match value exceeding the outside option. A "lean-out" general equilibrium in which $\rho_i = 0$ results in the monopsonist taking all of the surplus, that is, the monopsonist will optimally set the initial wage offer equal to the outside option, and no employee types will negotiate. By contrast, Proposition 4 finds that a lean-in general equilibrium with $\rho_i = 1$ leads to employees taking all of the surplus. Next, consider a competitive market in which there exist at least two employers with similar match values that are typically in head-to-head competition for employees (i.e., the distribution F places high probability mass on both employers being present in an employee's portfolio). Different general equilibria corresponding to different negotiation frequencies yield nearly identical outcomes, as the employers compete away surplus even without the pressure of employee-induced negotiations, similar to competition in Bertrand's model.

7 Conclusions

Salary negotiations are widespread in some labor markets and have the potential to affect not only individual job choices and amenities, but also market level pay equity and match efficiency. In this paper, we set out to bridge the gap between canonical full-information models of bargaining with the realities we observe in the field. By leveraging a field experiment in the US technology sector, we uncover that uncertainty about offer negotiability (even more so than the uncertainty of *how* to negotiate) prevents employees from fully exercising their bargaining power.

Our empirical findings reveal that when employees are provided with information that encourages negotiation, rates of negotiation rise and so too does average compensation. This is especially true among women, who make up a minority of US tech employees.

Based on these empirical insights, we model negotiation choices over a portfolio of offers, allowing for uncertainty about the (potentially negative) response of the employer. This model predicts that employees negotiate their best offers, making counterproposals that exceed their outside option. Indeed this prediction bears out in the observational data: compared to initial offers, final offers are more dispersed on average, rather than more equal, as predicted by Bertrand competition.

Our model allows us to consider policies that affect negotiation rates at the market level, allowing employers to strategically respond by adjusting their stance toward negotiations and also their initial wage offers. We compare the status quo with the lean-in negotiation equilibrium where negotiations are encouraged by a self-enforcing messaging campaign mimicking a scaled up version of our experimental encouragement treatment. We find the lean-in equilibrium maximizes employee welfare, efficiency, and equity. Moreover, our model suggests that moving to a lean-in equilibrium has the largest benefits in imperfectly competitive labor markets.

In summary, we believe that our study offers a unified framework that explains why negotiation behavior varies across job candidates for the same job, and highlights the potential gains from promoting negotiation.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Agan, A., B. Cowgill, and L. K. Gee (2020). Do workers comply with salary history bans? a survey on voluntary disclosure, adverse selection, and unraveling. *AEA Papers and Proceedings 110*, 215–219.

Akerlof, G. A. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

Ausubel, L. M., P. Cramton, and R. J. Deneckere (2002). Chapter 50 bargaining with incomplete information. Volume 3 of *Handbook of Game Theory with Economic Applications*, pp. 1897–1945. Elsevier.

Babcock, L. and S. Laschever (2003). *Women Don't Ask: Negotiation and the Gender Divide*. Princeton University Press.

Backus, M., T. Blake, B. Larsen, and S. Tadelis (2020). Sequential bargaining in the field: Evidence from millions of online bargaining interactions. *The Quarterly Journal of Economics 135*(3), 1319–1361.

Becker, G. M., M. H. Degroot, and J. Marschak (1963). Stochastic models of choice behavior. *Behavioral Science 8*(1), 41–55.

Biasi, B. and H. Sarsons (2021). Flexible wages, bargaining, and the gender gap. *The Quarterly Journal of Economics 137*(1), 215–266.

Business Insider (2025). The AI coding apocalypse. *February 25, 2025*.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006). Wage Bargaining with On-the-Job Search: Theory and Evidence. *Econometrica 74*(2), 323–364.

Caldwell, S., I. Haegele, and J. Heining (2024). Bargaining and inequality in the labor market. *Working Paper*.

Capozza, F. (2024). Beliefs about the gender gap in salary negotiations. *Working Paper*.

Card, D., A. R. Cardoso, and P. Kline (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics 131*(2), 633–686.

Cason, T. N. and C. R. Plott (2014). Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. *Journal of Political Economy 122*(6), 1235–1270.

Chade, H. and L. Smith (2006). Simultaneous Search. *Econometrica 74*(5), 1293–1307.

Chatterjee, K. and W. Samuelson (1983). Bargaining under Incomplete Information. *Operations Research 31*(5), 835–851.

Cortes, P., J. French, J. Pan, and B. Zafar (2024). Gender differences in negotiations and labor market outcomes: Evidence from an information intervention with college students. *NBER Working Paper No. 32154*.

Cullen, Z. (2024). Is pay transparency good? *Journal of Economic Perspectives 38*(1), 153–80.

Cullen, Z., S. Li, and R. Perez-Truglia (2022). What's my employee worth? the effects of salary benchmarking. *Review of Economic Studies, forthcoming*.

Cullen, Z. and B. Pakzad-Hurson (2023). Equilibrium Effects of Pay Transparency. *Econometrica 91*(3), 765–802.

Cullen, Z., B. Pakzad-Hurson, and R. Perez-Truglia (2025). Home sweet home: How much do employees value remote work? *AEA Papers and Proceedings 115*, 276–81.

Cullen, Z. and R. Perez-Truglia (2022). How Much Does Your Boss Make? The Effects of Salary Comparisons. *Journal of Political Economy 130*(3), 766–822.

Dannals, J. E., J. J. Zlatev, N. Halevy, and M. A. Neale (2021). The dynamics of gender and alternatives in negotiation. *Journal of Applied Psychology 106*(11), 1655–1672.

Deloitte (2022). Rise of women in tech leadership. Accessed: 2025-05-27.

Doval, L. (2018). Whether or not to open pandora's box. *Journal of Economic Theory 175*, 127–158.

Enwemeka, Z. (2016). To close the wage gap, boston hopes salary negotiation workshops will create a culture shift. *WBUR, August 26, 2016*.

Exley, C. L., M. Niederle, and L. Vesterlund (2020). Knowing when to ask: The cost of leaning in. *Journal of Political Economy 128*(3), 816–854.

Fidelity (2022). 85 percent of americans who negotiated a job offer were successful. *Fidelity Smart Money, https://www.fidelity.com/learning-center/smart-money/how-to-negotiate-salary*.

Fuchs, W. and A. Skrzypacz (2010). Bargaining with Arrival of New Traders. *American Economic Review 100*(3), 802–836.

Fudenberg, D., D. Levine, and J. Tirole (1985). Infinite-horizon models of bargaining with one-sided incomplete information. In A. E. Roth (Ed.), *Game-Theoretic Models of Bargaining*, Chapter 5, pp. 73–98. Cambridge University Press.

Gentile Passaro, D., F. Kojima, and B. Pakzad-Hurson (2024). Equal pay for *Similar* work. *arXiv preprint arXiv:2306.17111*.

Hall, R. E. and A. B. Krueger (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics 4*(4), 56–67.

Kray, L. (2015). The best way to eliminate the gender pay gap? Ban salary negotiations. *Washington Post, May 21, 2015*.

Kray, L. J., J. A. Kennedy, and M. Lee (2024). Now, women do ask: A call to update beliefs about the gender pay gap. *Academy of Management Discoveries 10*(1), 11–37.

Lachowska, M., A. Mas, R. Saggio, and S. A. Woodbury (2022). Wage posting or wage bargaining? A test using dual jobholders. *Journal of Labor Economics 40*(51).

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies 61*(3), 397–415.

Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*(2), 265–281.

Nash, J. F. (1950). The Bargaining Problem. *Econometrica 18*(2), 155–162.

Perez-Truglia, R. (2023). The effects of pay transparency: A brief review. *Kenan Institute's Grand Challenge Report*.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica 70*(6), 2295–2350.

Recalde, M. P. and L. Vesterlund (2023). Gender differences in negotiation: Can interventions reduce the gap? *Annual Review of Economics 15*, 633–657.

Reny, P. J. (2020). Nash Equilibrium in Discontinuous Games. *Annual Review of Economics 12*, 439–470.

Roussille, N. (2024). The central role of the ask gap in gender pay inequality. *The Quarterly Journal of Economics 139*(3), 1557–1610.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica 50*(1), 97–109.

Sandberg, S. and N. Scovell (2013). *Lean In: Women, Work, and the Will to Lead*. New York: Alfred A. Knopf.

Sinclair, M., J. O'Toole, M. Malawaraarachchi, and K. Leder (2012). Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Medical Research Methodology 12*(1), 132.

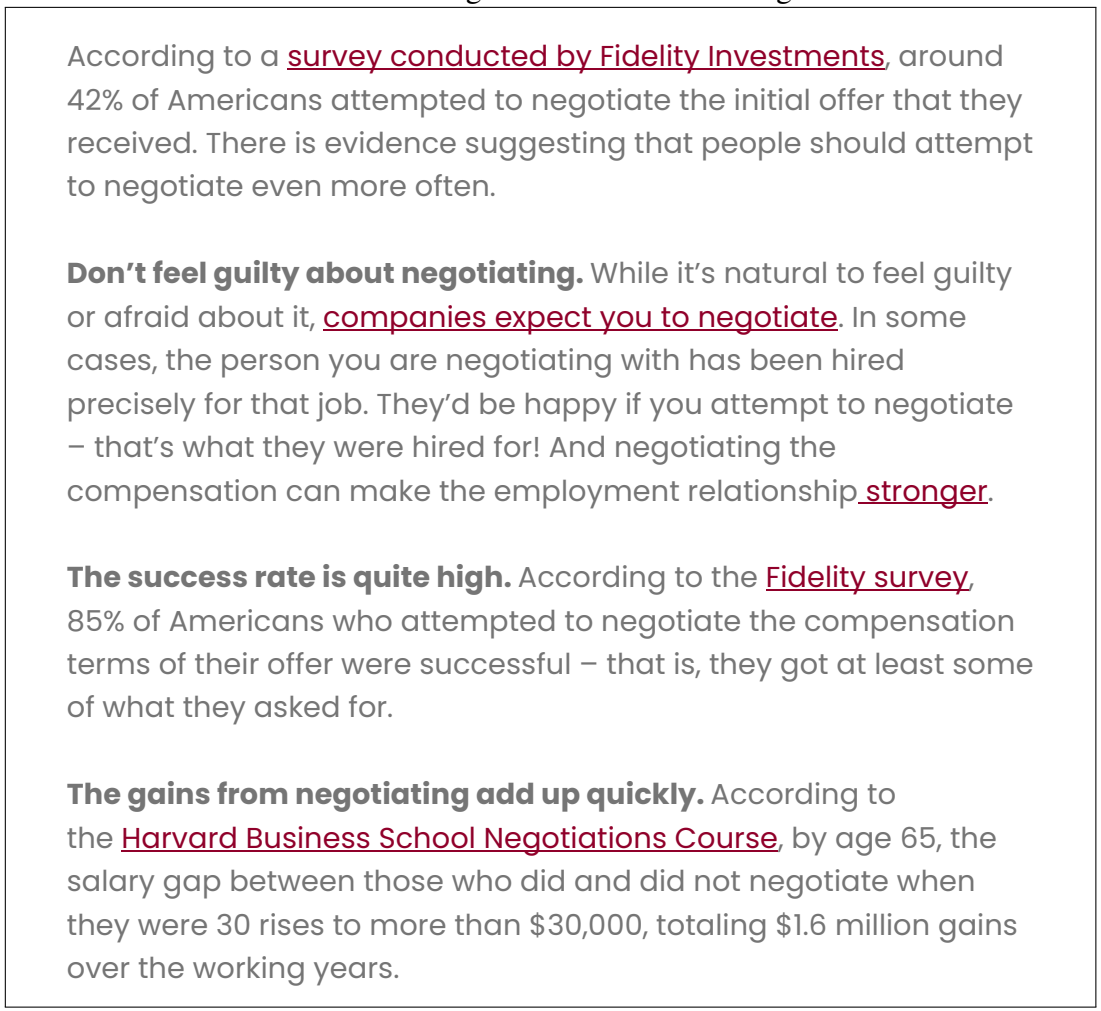
Sockin, J. (2021). Show Me the Amenity: Are Higher-Paying Firms Better All Around? *CESifo Working Paper No. 9842*.

Stole, L. A. and J. Zwiebel (1996). Intra-firm bargaining under non-binding contracts. *Review of Economic Studies 63*(3), 375–410.

Weitzman, M. L. (1979). Optimal Search for the Best Alternative. *Econometrica 47*(3), 641–654.

Figure 1: Treatment Messages

PANEL A: Negotiation Lessons Message



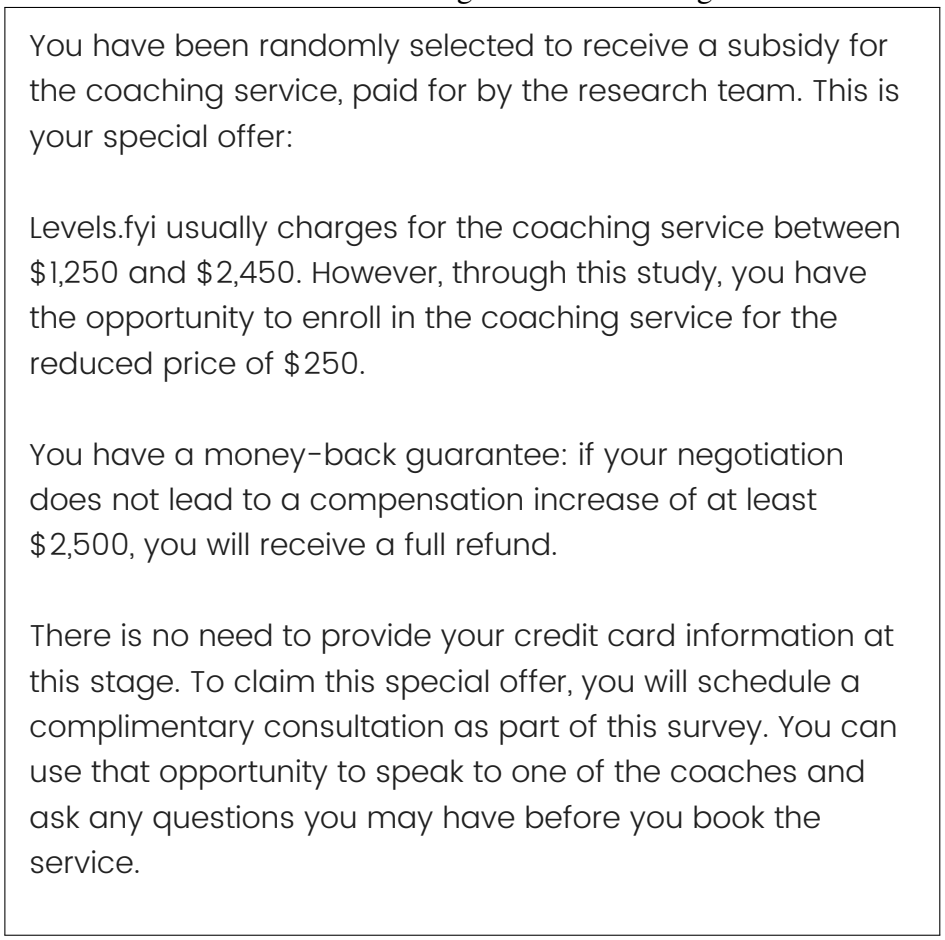
According to a survey conducted by Fidelity Investments, around 42% of Americans attempted to negotiate the initial offer that they received. There is evidence suggesting that people should attempt to negotiate even more often.

Don't feel guilty about negotiating. While it's natural to feel guilty or afraid about it, companies expect you to negotiate. In some cases, the person you are negotiating with has been hired precisely for that job. They'd be happy if you attempt to negotiate – that's what they were hired for! And negotiating the compensation can make the employment relationship stronger.

The success rate is quite high. According to the Fidelity survey, 85% of Americans who attempted to negotiate the compensation terms of their offer were successful – that is, they got at least some of what they asked for.

The gains from negotiating add up quickly. According to the Harvard Business School Negotiations Course, by age 65, the salary gap between those who did and did not negotiate when they were 30 rises to more than $30,000, totaling $1.6 million gains over the working years.

PANEL B: Coaching Treatment Message



You have been randomly selected to receive a subsidy for the coaching service, paid for by the research team. This is your special offer:

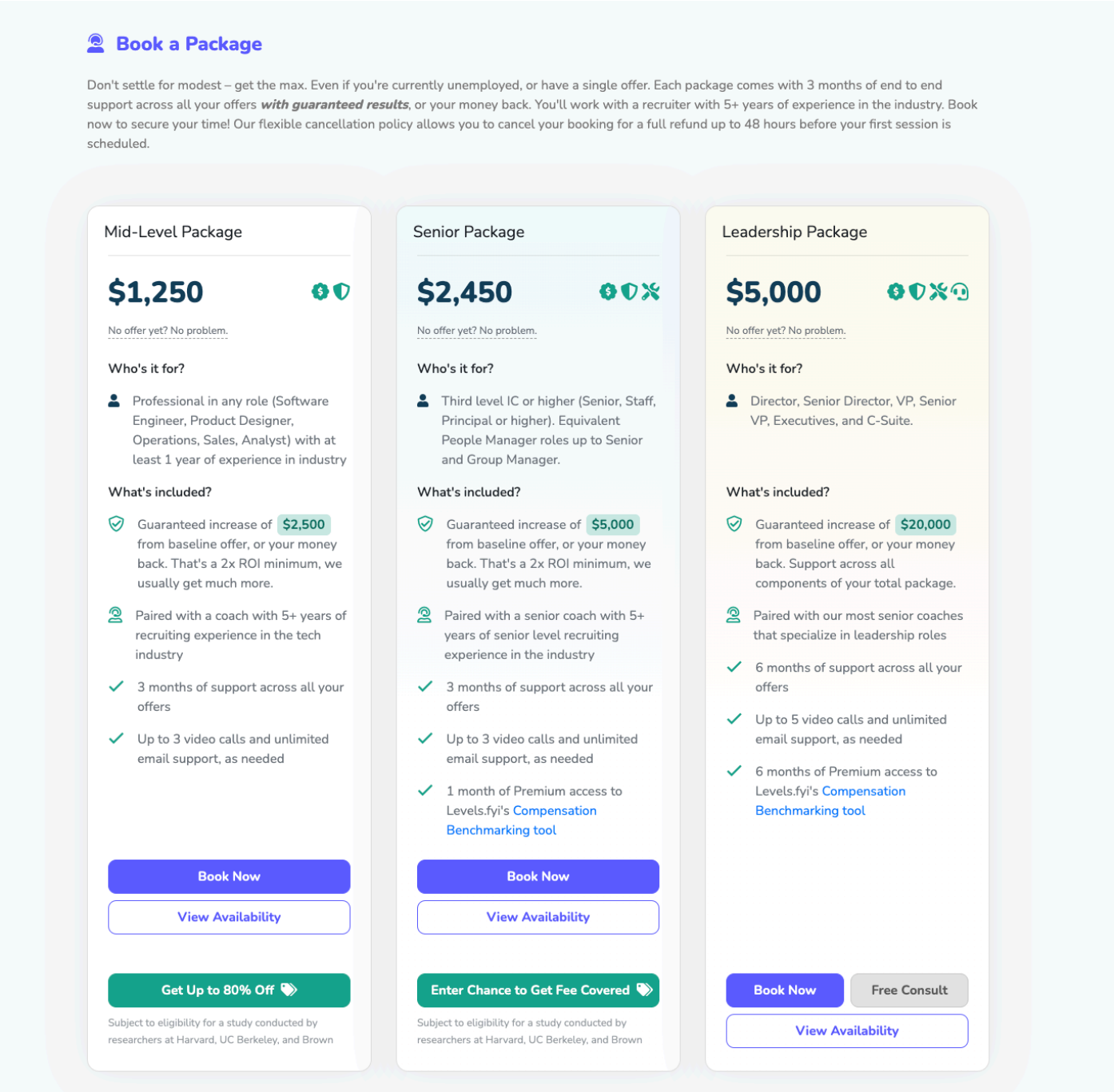
Levels.fyi usually charges for the coaching service between $1,250 and $2,450. However, through this study, you have the opportunity to enroll in the coaching service for the reduced price of $250.

You have a money-back guarantee: if your negotiation does not lead to a compensation increase of at least $2,500, you will receive a full refund.

There is no need to provide your credit card information at this stage. To claim this special offer, you will schedule a complimentary consultation as part of this survey. You can use that opportunity to speak to one of the coaches and ask any questions you may have before you book the service.

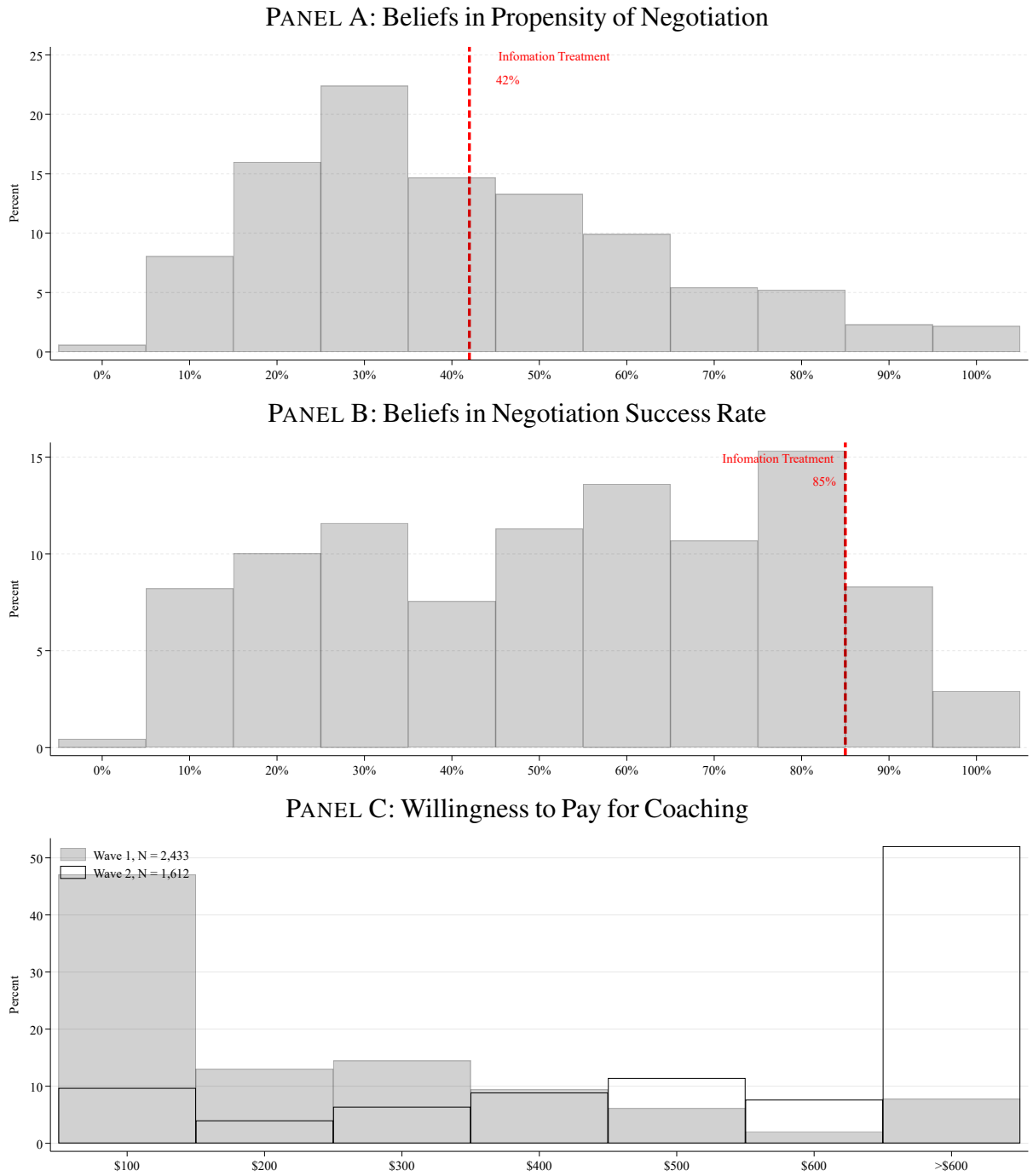
Notes: Panel A provides a screenshot of the additional message shown to subjects in the encouragement treatment, while Panel B shows the screenshot of the addi shown to subjects in the coaching treatment.

Figure 2: Screenshot of the Website with Survey Links



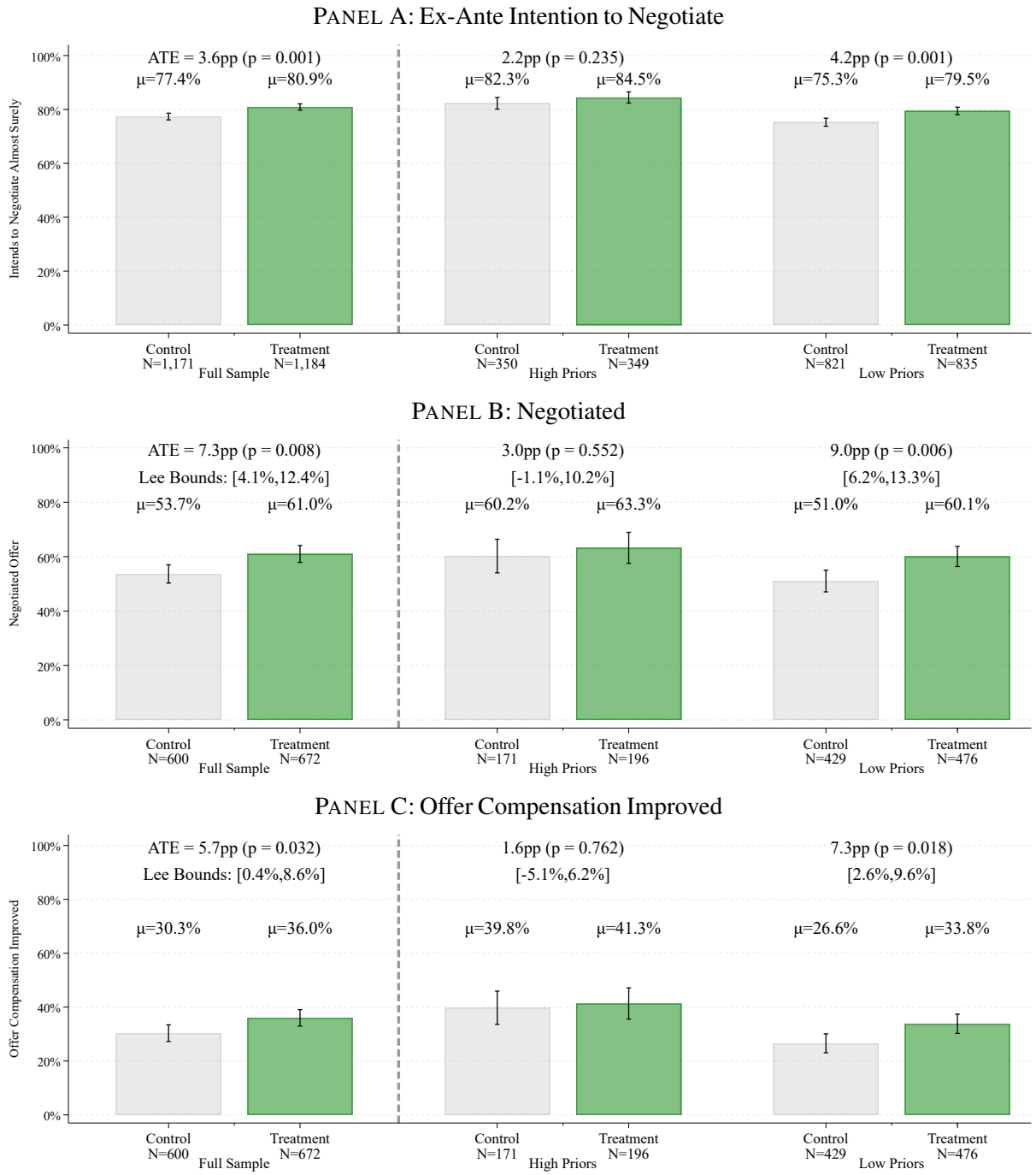
Notes: This is a screenshot of the website with the links to the baseline survey.

Figure 3: Distribution of Prior Beliefs and Willingness to Pay



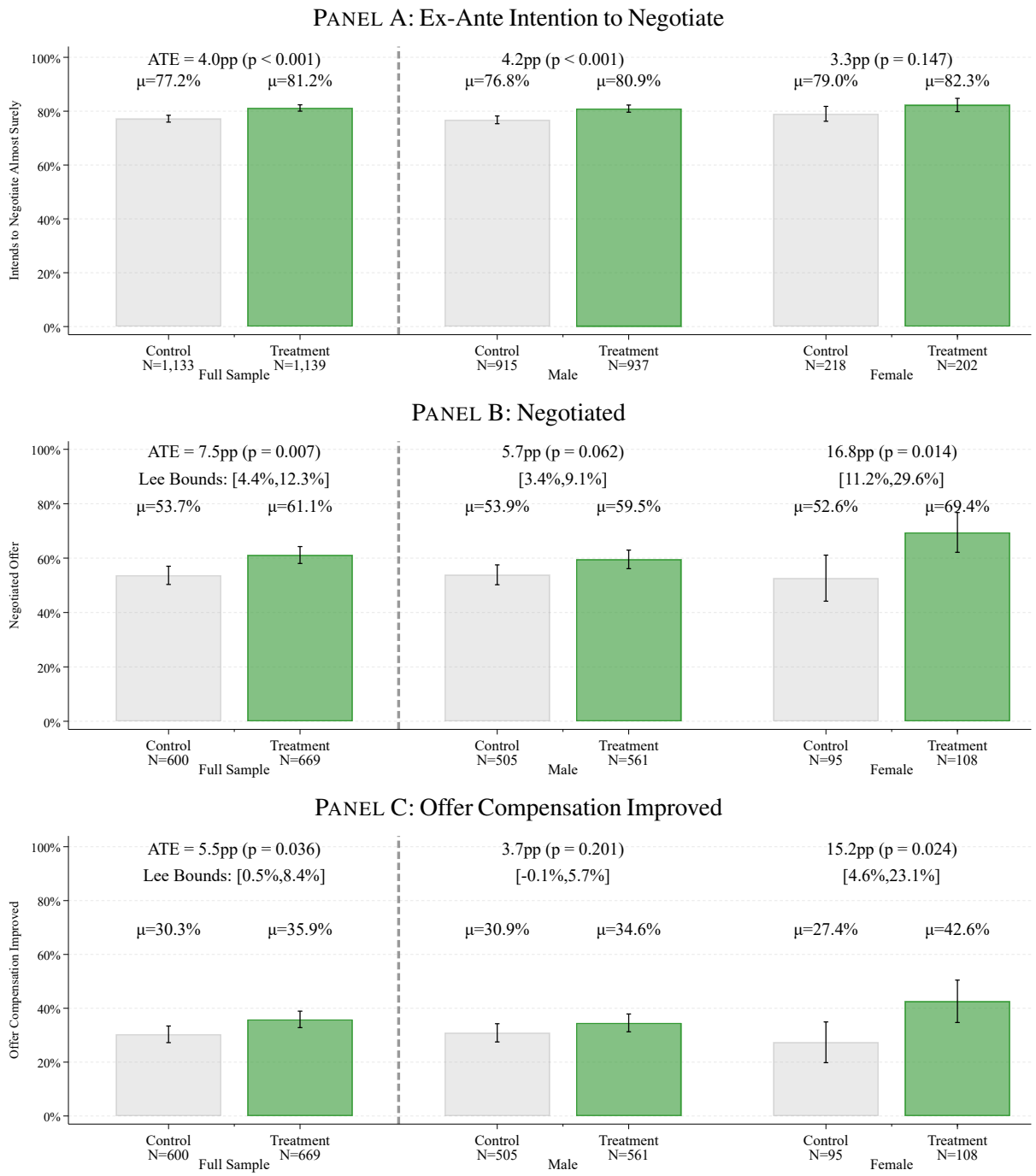
Notes: Panel A is based on the question "What percent of people do you think try to negotiate the compensation terms after receiving a job offer?" Panel B is based on the question "Among individuals who attempt to negotiate, what percentage of them do you think were successful? In other words, what percentage got at least some of what they asked for?" The dashed lines show the numbers we presented on the encouragement lesson screen. Panel C includes all valid responses to the willingness-to-pay questions, representing 97% of the full sample. We exclude inconsistent responses in which participants chose the coaching service over a higher gift card amount but not over a lower amount.

Figure 4: The Effects of Encouragement Treatment and Heterogeneity by Priors (Wave 1)



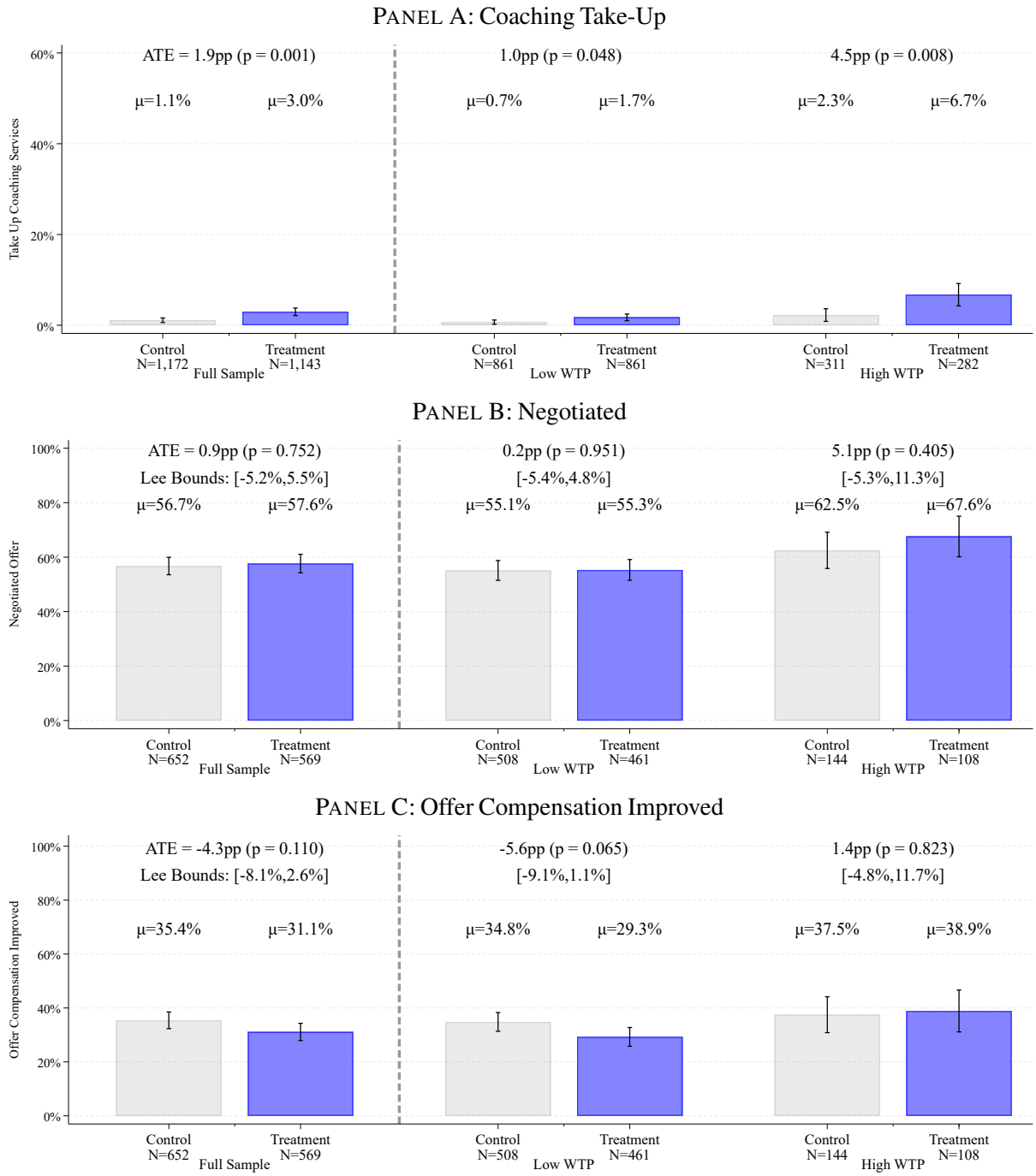
Notes: Main outcomes by encouragement treatment status (with the treated group in green bars and 90% standard errors in brackets). Each panel corresponds to a different dependent variable: Panel A for ex-ante intention to negotiate, Panel B for negotiating an offer, and Panel C for having offer compensation increased. In each panel, the first two bars represent the full wave 1 sample, and the right four bars split the sample based on prior beliefs. Participants are categorized as "Low Priors" if their prior estimates about negotiation rates and success rates were lower than the provided information.

Figure 5: The Effects of Encouragement Treatment and Heterogeneity by Gender (Wave 1)



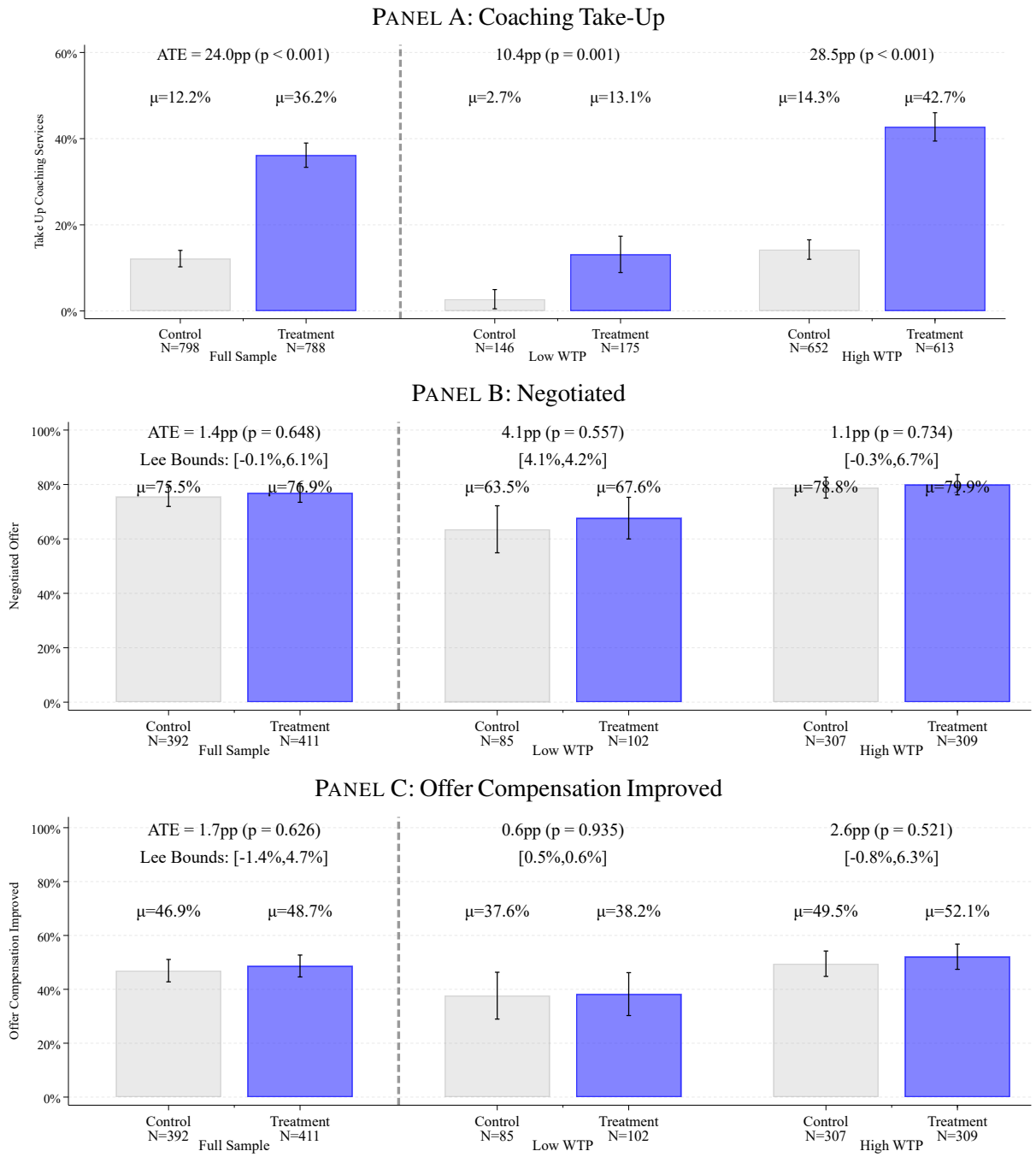
Notes: Main outcomes by encouragement treatment status (with the treated group in green bars and 90% standard errors in brackets). Each panel corresponds to a different dependent variable: Panel A for ex-ante intention to negotiate, Panel B for negotiating an offer, and Panel C for having offer compensation increased. In each panel, the first two bars represent the full wave 1 sample, and the right four bars split the sample based on gender.

Figure 6: The Effects of Coaching Treatment and Heterogeneity by Stated Willingness to Pay (Wave 1)



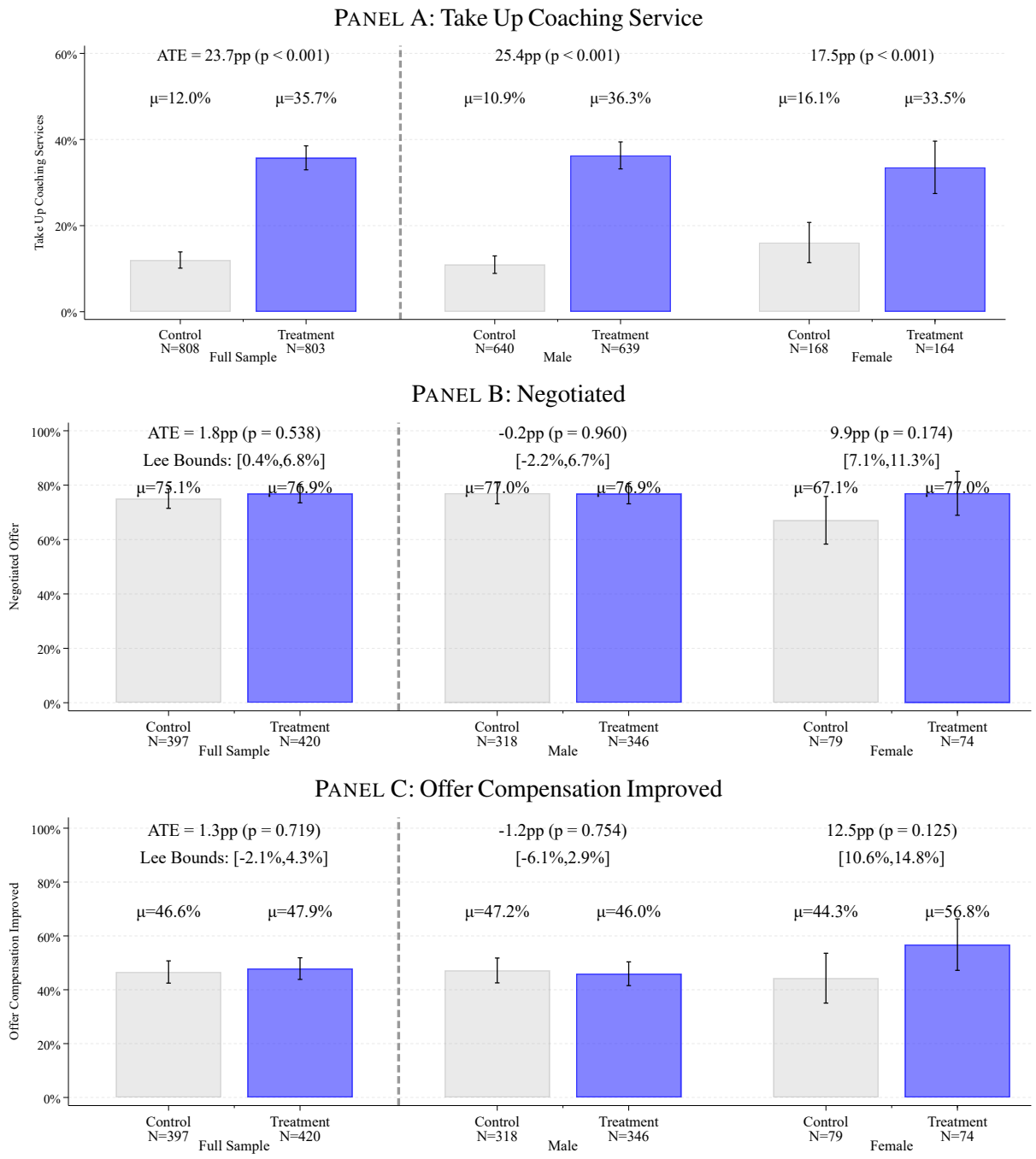
Notes: Main outcomes by coaching treatment status (with the treated group in blue bars and 90% standard errors in brackets). Each panel corresponds to a different dependent variable: Panel A for taking up coaching services, Panel B for negotiating an offer, and Panel C for having offer compensation increased. In each panel, the first two bars represent the full wave 1 sample, and the right four bars split the sample based on their stated WTP for coaching services. Participants are categorized as "Low WTP" if their stated WTP is lower than $300.

Figure 7: The Effects of Coaching Treatment and Heterogeneity by Stated Willingness to Pay (Wave 2)



Notes: Main outcomes by coaching treatment status (with the treated group in blue bars and 90% standard errors in brackets). Each panel corresponds to a different dependent variable: Panel A for taking up coaching services, Panel B for negotiating an offer, and Panel C for having offer compensation increased. In each panel, the first two bars represent the full wave 2 sample, and the right four bars split the sample based on stated WTP for coaching services. Participants are categorized as "Low WTP" if their stated WTP is lower than $300.

Figure 8: The Effects of Coaching Treatment and Heterogeneity by Gender (Wave 2)



Notes: Main outcomes by coaching treatment status (with the treated group in blue bars and 90% standard errors in brackets). Each panel corresponds to a different dependent variable: Panel A for taking up coaching services, Panel B for negotiating an offer, and Panel C for having offer compensation increased. In each panel, the first two bars represent the full wave 2 sample, and the right four bars split the sample based on gender.

A Proofs, and Additional Theoretical Results

Proof of Proposition 1:

Proof. It suffices to consider only employers with match values greater than the employee's outside option, so without loss of generality, let $v_1 > v_2 > ... > v_N > 0$. Because the employee's match value is lower at employer N than at all other employers, the employee accepts a job at employer N only if, for each employer $i < N$, she was either rejected from i, or she did not make a counteroffer to employer i.

Consider the following algorithm.

Stage [1] Let n be the smallest-index employer such that either $\hat{\rho}_n = 1$ or $w_n = v_n$. If no such n exists, then define $n = N+1$. If $n < N+1$, then $o_n^* = v_n - w_n$, and $o_i^* = 0$ for all $i > n$. If $n = 1$, then terminate the algorithm. Otherwise, go to Stage [2].

The remainder of this algorithm considers $i < n$. We inductively construct the optimal counteroffers via a double induction argument.

Stage [2] Let $\theta_{n-1} := v_n \mathbb{1}_{n<N+1}$.

Stage [2.1] If $\hat{\rho}_{n-1} v_{n-1} + (1-\hat{\rho}_{n-1})\theta_{n-1} \geqslant w_{n-1}$, then set $o_{n-1}^{[2.1]} = v_{n-1} - w_{n-1}$ and set $\theta_{n-2}^{[2.1]} = \hat{\rho}_{n-2} v_{n-2} + (1-\hat{\rho}_{n-2})\theta_{n-1}$. Otherwise, set $o_{n-1}^{[2.1]} = 0$ and set $\theta_{n-2}^{[2.1]} = w_{n-1}$. If $n-1 = 1$ then terminate the algorithm and let $o_i^{[2.1]} = o_i^*$ for all $i \leqslant n$. Otherwise, go to Stage [2.2].

Stage [2.k] Let $o_i^{[2.k]} = o_i^{[2.k-1]}$ for all $i \in \{n-(k-1), ..., n\}$. If $\hat{\rho}_{n-k} v_{n-k} + (1-\hat{\rho}_{n-k})\theta_{n-k}^{[2.k-1]} \geqslant w_{n-k}$, then set $o_{n-k}^{[2.k]} = v_{n-k} - w_{n-k}$ and set $\theta_{n-k-1}^{[2.k]} = \hat{\rho}_{n-k-1} v_{n-k-1} + (1-\hat{\rho}_{n-k-1})\theta_{n-k}^{[2.k-1]}$, moreover, if $n-k = 1$ then terminate the algorithm and let $o_i^{[2.k]} = o_i^*$ for all $i \leqslant n$ and if $n-k > 1$ then go to Stage [2.k+1]. If $\hat{\rho}_{n-k} v_{n-k} + (1-\hat{\rho}_{n-k})\theta_{n-k}^{[2.k-1]} < w_{n-k}$, go to Stage [3].

Stage [m] Let $K^{[m]}$ be the final employer considered in Stage $[m-1]$. Let $o_{K^{[m]}}^{[m.1]} = 0$, let $o_i^{[m.1]} = 0$ for all $i > K^{[m]}$ such that $v_i \leqslant w_{K^{[m]}}$ or $\hat{\rho}_i = 0$, and let $o_j^{[m.1]} = v_j - w_j$ for all other employers $j > K^{[m]}$. Let $L^{[m]}$ be the set of employers such that for each $\ell \in L^{[m]}$, $w_{K^{[m]}} < v_\ell < v_{K^{[m]}}$ and $\hat{\rho}_\ell > 0$. Set $\theta_{K^{[m]}-1}^{[m.1]} = \sum_{j=1}^{|L^{[m]}|} \left(\prod_{\ell=1}^{j-1} (1-\hat{\rho}_{k+\ell}) \right) \hat{\rho}_{k+j} v_{k+j} + \left(\prod_{\ell=1}^{|L^{[m]}|} (1-\hat{\rho}_{k+\ell}) \right) w_{K^{[m]}}$.

Stage $[m.1\ ]$ If $K^{[m]}-1=0$ then terminate the algorithm and let $o_i^{[m.1]}=o_i^*$ for all $i\leqslant n$. Otherwise, if $\hat{\rho}_{K^{[m]}-1}v_{K^{[m]}-1}+(1-\hat{\rho}_{K^{[m]}-1})\theta^{[m.1]}_{K^{[m]}-1}\geqslant w_{K^{[m]}-1}$, then set $o^{m.1}_{K^{[m]}-1}=v_{K^{[m]}-1}-w_{K^{[m]}-1}$ and set $\theta^{m.1}_{K^{[m]}-2}=\hat{\rho}_{K^{[m]}-2}v_{K^{[m]}-2}+(1-\hat{\rho}_{K^{[m]}-2})\theta^{m.1}_{K^{[m]}-1}$, then go to Stage $[m.2]$. If $\hat{\rho}_{K^{[m]}-1}v_{K^{[m]}-1}+(1-\hat{\rho}_{K^{[m]}-1})\theta^{[m.1-1]}_{K^{[m]}-1}<w_{K^{[m]}-1}$, go to Stage $[m+1]$.

Stage $[m.k\ ]$ Let $o_i^{[m.k]}=o_i^{[m.k-1]}$ for all $i\in\{K^{[m]}-(k-1),...,n\}$. If $\hat{\rho}_{K^{[m]}-k}v_{K^{[m]}-k}+(1-\hat{\rho}_{K^{[m]}-k})\theta^{[m.k-1]}_{K^{[m]}-k}\geqslant w_{K^{[m]}-k}$, then set $o^{[m.k]}_{K^{[m]}-k}=v_{K^{[m]}-k}-w_{K^{[m]}-k}$ and set $\theta^{[m.k]}_{K^{[m]}-k-1}=\hat{\rho}_{K^{[m]}-k-1}v_{K^{[m]}-k-1}+(1-\hat{\rho}_{K^{[m]}-k-1})\theta^{[m.k-1]}_{K^{[m]}-k}$, moreover, if $K^{[m]}-k=0$ then terminate the algorithm and let $o_i^{[m.k]}=o_i^*$ for all $i\leqslant n$ and if $K^{[m]}-k>0$ then go to Stage $[m.k+1]$. If $\hat{\rho}_{K^{[m]}-k}v_{K^{[m]}-k}+(1-\hat{\rho}_{K^{[m]}-k})\theta^{[m.k-1]}_{K^{[m]}-k}<w_{K^{[m]}-k}$, go to Stage $[m+1]$.

We claim that the algorithm terminates, does so at an optimal portfolio of counteroffers, and that this optimal portfolio is generically unique.

To see that the algorithm terminates, suppose for contradiction that it does not. This means that at no stage $[m]$, $m>2$ does the algorithm reach a substage $[m.k]$ such that $K^{[m]}-k=0$ and $\hat{\rho}_{K^{[m]}-k}v_{K^{[m]}-k}+(1-\hat{\rho}_{K^{[m]}-k})\theta^{[m.k-1]}_{K^{[m]}-k}\geqslant w_{K^{[m]}-k}$. However, by construction, $(K^{[m]})_{m>2}$ is a strictly decreasing sequence of integers, i.e., a lower index employer is the first employer referenced at the beginning of each Stage. Therefore, there is some stage $[m^*]$, $m\leqslant n+1$ such that $K^{[m^*]}=1$. Then the algorithm immediately terminates, i.e., by Stage $[m^*.1]$. This contradicts that the algorithm never terminates, which completes the desired argument.

We now claim that the outcome of the algorithm o^* is optimal, and uniquely so generically.

To see this, let o' be an optimal counteroffer portfolio. First note that, defining n as in Stage $[1]$, it must be the case that $o'_n=v_n-w_n$ and $o'_i=0$ for all $i>n$ in any optimal counteroffer portfolio. Therefore, $o'_j=o^*_j$ for all $j\geqslant n$.

Consider any Stage $[m-1]$, $m-1\geqslant 2$. We claim that if the algorithm terminates at the conclusion of Stage $[m-1]$ then o^* is generically uniquely optimal, and if it does not terminate, then $(o^{[m.1]}_{K^{[m]}},o^{[m.1]}_{K^{[m]+1}},...,o^{[m.1]}_N)$ is generically the uniquely optimal portfolio of counteroffers supposing the employee's initial portfolio comprised of only offers from employers $(K^{[m]},K^{[m]+1},...,N)$. Our argument is by induction on the index of Stages.

Base case, $m-1=2$**:** If the algorithm terminates in Stage $[2]$ then, by construction, the employee optimally negotiates with all employers $j<n-1$, and the employee has strict incentives

to negotiate with all such employers if $\hat{\rho}_{n-k}v_{n-k}+(1-\hat{\rho}_{n-k})\theta_{n-k}^{[2.k-1]} > w_{n-k}$ in all Stages $[2.k]$, $k>1$ which is generically satisfied given that $\hat{\rho}_{n-k}v_{n-k}+(1-\hat{\rho}_{n-k})\theta_{n-k}^{[2.k-1]} \geqslant w_{n-k}$ is satisfied (note that by a similar argument, the employee's generically optimal choice is defined for employer $n-1$). If the algorithm does not terminate in Stage [2] then there is some final substage $[2.\hat{k}]$, in which the employee considers negotiation with employer $K^{[3]}$. At the conclusion of stage $[2.\hat{k}]$, the (generically unique) optimal negotiation decision has been determined, by the most recent argument, for all employers $i > K^{[3]}$ under the inductive hypothesis that the employee negotiates with all employers $i \geqslant K^{[3]}$, however, this assumption has been violated by the fact that the transition to Stage [3] implies that the employer does not optimally negotiate with employer $K^{[3]}$ given an expected value $\theta_{K^{[3]}}^{[2.\hat{k}]}$ conditional on being rejected from employer $j \leqslant K^{[3]}$ because

$$\hat{\rho}_{K^{[3]}}v_{K^{[3]}}+(1-\hat{\rho}_{K^{[3]}})\theta_{K^{[3]}}^{[2.\hat{k}]} < w_{K^{[3]}}. \tag{A.1}$$

Assume, subject to later verification, that in any optimal vector of counteroffers given a portfolio comprised only of offers from employers $(K^{[3]},K^{[3]+1},...,N)$, that the employer elects not to negotiate with employer $K^{[3]}$. Then, given this portfolio, the employee achieves a payoff of at least $w_{K^{[3]}}$ and must negotiate with employer $i > K^{[3]}$ if and only if $v_i \geqslant w_{K^{[3]}}$, $\hat{\rho}_i > 0$, and $w_i < v_i$, where by genericity we assume that there are no employers i such that $v_i = w_{K^{[3]}}$. Doing so implies that, the employee matches with employer $K^{[3]}$ if and only if she is rejected from all such employers i. This mechanically lowers the employee's continuation value, conditional on being rejected from all such employers i and $K^{[3]}$ to a value lower than $\theta_{K^{[3]}}^{[2.\hat{k}]}$. Therefore, our assumption subject to verification, and the proof of the Base Case, is shown to be correct if and only if the employee optimally elects not to negotiate with employer $K^{[3]}$ given this lower continuation value. But this is satisfied, because if Equation (A.1) is satisfied, it is also satisfied if instead the employee's continuation value $\theta_{K^{[3]}}^{[2.\hat{k}]}$ is reduced.

Induction case, $m-1>2$: Our argument thus far implies that $(o_{K^{[m]}}^{[m.1]},o_{K^{[m]+1}}^{[m.1]},...,o_{N}^{[m.1]})$ is generically the uniquely optimal portfolio of counteroffers supposing the employee's initial portfolio comprised of only offers from employers $(K^{[m]},K^{[m]+1},...,N)$. This is clearly true if $K^{[m]}-1=0$, i.e $K^{[m]}=1$. Otherwise, the construction of $\theta_{K^{[m]}-1}^{[m.1]}$ gives the employee's optimal continuation

value given that she is rejected from all employers j such that $j < K^{[m]}$, or $v_j \geqslant w_{K^{[3]}}$, $\hat{\rho}_j > 0$, and $w_j < v_i$. The remainder of the argument follows similar logic as that of the Base Case.

We have already shown that the algorithm terminates at some Stage $[m]$, $m \leqslant n+1$. When it does so, by construction, it specifies a value o_i^* for each employer i. By the inductive argument presented above, it must be the case that o_i^* is generically the unique negotiation vector.

□

Proof of Proposition 2

Proof of Part 1: Consider an employer ℓ such that $v_\ell > v_j$. We argue that the employee makes a counteroffer to employer ℓ. To this end, it is either the case that $w_\ell > w_j$ or $w_\ell \leqslant w_j$:

- Suppose $w_\ell > w_j$. Then it must be the case that $o_\ell^* > 0$, as desired, or else j would not be the highest-initial-offer employer that the employee does not make a counteroffer to.
- Suppose $w_\ell \leqslant w_j$, and further suppose for contradiction that $o_\ell^* = 0$. First, note that there is a strictly positive probability that the employee is matched with employer j; if not, then because we assume $\hat{\rho}_i \in (0,1)$ for all i, there exists some employer i' with $v_{i'} = w_{i'}$ such that $v_{i'} > v_j$. But combining the inequality and equality in the previous sentence implies that $w_{i'} > w_j$, which contradicts the ongoing assumption that j is the highest-initial-offer employer for which $o_j^* = 0$. Given the strictly positive probability of matching with employer j, we claim that $w_\ell \leqslant w_j$ implies that the employee would be better off making a counteroffer to employer ℓ. This claim is clear if $w_\ell < w_j$, and if $w_\ell = w_j$, then the maintained assumption that $v_\ell > v_j$ implies that $v_\ell - w_\ell > 0$ meaning that there are strictly positive gains from negotiating with employer ℓ. Therefore, $o_\ell^* = 0$ contradicts the assumed optimality of the counteroffers.

In the two exhaustive cases above, we arrive at a contradiction if $o_j^* = 0$, thus completing the argument. □

Proof of Part 2: This follows from inspection of the algorithm provided in the proof of Proposition 1—at each stage, the employee negotiates with the indicated employer i if and only if the stated condition is satisfied. □

Proof of Proposition 3

Proof of Parts 1 and 2. The first two parts of this proposition follow straightforwardly. To see that a partial equilibrium exists, note that we have already established in Proposition 1 that for any portfolio of offers and beliefs $\{w_i, v_i, \hat{\rho}_i\}_{i \in N}$ there exists an optimal portfolio of counteroffers for the selected employee.[64] Therefore, any given strategy profile of employers generates a distribution over payoffs for each employer; it follows from established equilibrium-existence results (see, e.g., Reny, 2020, Theorem 3) that the truncated game which "automates" the portfolio bargaining problem and directly yields each employer the expected payoff given the selected strategies by all employers contains a Nash equilibrium.[65] Our game, which is a concatenation of the employer-side decisions followed by the employee-side decisions, therefore always has a partial equilibrium, as rational expectations are not required. To see that a general equilibrium exists if the support of F contains only the "correct" beliefs $\hat{\rho}$, we claim that each employer i's payoff is unaffected by the choice of ρ_i. This follows for two reasons: first, as can be seen in Part 2 of Proposition 2, each employee type's decision to make a counteroffer is affected by the belief $\hat{\rho}_i$ but not by the choice ρ_i. Second, the payoff of any employer is zero if the employee negotiates with it, as either the employee receives the entire match value, or the employer fails to hire the employee. Therefore, the same argument ensuring the existence of a partial equilibrium in any labor market implies that, for a distribution F with a single belief vector, there is a general equilibrium of the employer-side truncated game in which each employer selects $\rho_i = \hat{\rho}_i$. □

Proof of Part 3.

"if" direction Fix (N, Θ) and let $\hat{\rho}$ be a belief vector such that $\hat{\rho}_i \in \{0,1\}$ for all i.

First consider an employer $j \in N$ such that $\hat{\rho}_j = 0$, which implies that no employee type makes a counteroffer to j in general equilibrium. Because $w_j^g < v_j$ means by definition that j negotiates with no group-g employees, the claim is demonstrated if we can show that for any $g \in \{A, B\}$, $w_j^g < v_j$ is supported in general equilibrium in any labor market $(N, \Theta, F, \{v_i\}_{i \in N})$ with only belief vector $\hat{\rho}$ in the support of F. Holding fixed the strategies of other employers, consider two exhaustive cases. First, suppose that almost no employee type is matched to j for any $w_j^g < v_j$. By the full support assumption on F, and by Part 2 of Proposition 2, this implies that $v_j \leqslant \underline{\theta}$, so that any $w_j^g < v_j$ can be supported

[64] Proposition 1 assumed that the outside option $\theta = 0$ but this was merely a normalization that does not change the applicability to the current equilibrium setting.

[65] A similar "automation" argument used to show existence of equilibrium is presented in Gentile Passaro et al. (2024).

in general equilibrium. Second, suppose that at least one employee type is matched to j for some $w_j^g < v_j$. Then $v_j > \underline{\theta}$. Consider $\tilde{w}_j^g = \underline{\theta} + \epsilon$, where $\epsilon > 0$. For ϵ sufficiently small, the full support assumption on F and Part 2 of Proposition 2 imply that with positive probability, an employee type r will be selected such that: 1) $\theta^r < \tilde{w}_j^g$, 2) $g^r = g$, and 3) $N^r = \{j\}$. Because such an employee type r optimally does not make a counteroffer to j, r will be employed at employer j at wage $\tilde{w}_j^g < v_j$, and because such a type is selected with strictly positive probability, j earns positive expected profit by selecting initial wage offer $\tilde{w}_j^g$. Thus, j cannot set $w_j^g = v_j$ in general equilibrium, as this yields zero profit (from employees of group g). This completes the argument in the case that $\hat{\rho}_j = 0$.

Next, consider an employer $\ell \in \mathcal{N}$ such that $\hat{\rho}_\ell = 1$. For each employee type r of group g who interacts with subset of employers N including ℓ, the employee will match to employer ℓ if and only if $v_\ell > w_i^g$ for all employers $i \in N$ such that $\hat{\rho}_i = 0$ and $v_\ell > v_i$ for all employers $i \in N \setminus \{\ell\}$ such that $\hat{\rho}_i = 1$, regardless of w_ℓ^g. Therefore, for any labor market $(\mathcal{N}, \Theta, F, \{v_i\}_{i \in \mathcal{N}})$ with only belief vector $\hat{\rho}$ in the support of F, there is a general equilibrium in which employer ℓ sets $w_\ell^g = v_\ell$ for all $g \in \{A, B\}$. In such a general equilibrium, the negotiation probability at employer j is 1 by fiat, as desired.

"only if" direction The following argument demonstrates that $\hat{\rho}$ is not robustly self-enforcing if $\hat{\rho}_j \in (0,1)$ for some $j \in \mathcal{N}$.[66] Suppose for contradiction that given $(\mathcal{N}, \Theta)$, there exists a robustly self-enforcing belief vector $\hat{\rho}$ such that $\hat{\rho}_j \in (0,1)$ for some $j \in \mathcal{N}$. Consider a labor market $(\mathcal{N}, \Theta, F, \{v_i\}_{i \in \mathcal{N}})$ in which $v_j < \underline{\theta}$. Then no employee type make counteroffers, i.e., each employee optimally sets $o_i^* = 0$. Therefore, in order for the employer to negotiate with the proper probability $\hat{\rho}_i$, it must be that either $w_i^A = v_j$ or $w_i^B = v_j$, but not both: if $w_i^A < v_j$ and $w_i^B < v_j$, then $o_i^* = 0$ for all employee types implies that the negotiation probability at employer i is equal to $0 < \hat{\rho}_i$, and if $w_i^A = v_j = w_i^B$, then the negotiation probability at employer i is, by definition, equal to $1 > \hat{\rho}_i$. Similarly, the negotiation probability at employer i will equal $\hat{\rho}_i$ only if the probability of an employee being a member of some group $g \in \{A, B\}$ is equal to $\hat{\rho}_i$, and $w_i^g = v_i$ and $w_i^{-g} < v_i$ where $-g \neq g$. As the probability of an employee being a member of some group $g \in \{A, B\}$ being equal to $\hat{\rho}_i$ does not hold across all possible distributions F such that $\hat{\rho}$ is the unique belief vector in its support (indeed it fails to hold generically), $\hat{\rho}$ is not robustly self enforcing, leading to a contradiction. □

[66] We note the existence of, but do not present, more complicated arguments that yield the same conclusion.

Proof of Proposition 4

The argument behind this result follows from a similar one to that presented in the Proof of Proposition 5, and is therefore omitted.

Proof of Proposition 5

Proof. Let $\hat{\rho}$ be the only belief vector in the support of the labor market. We show that for any $\epsilon > 0$ there exists $\delta > 0$ such that if $\hat{\rho}_i > 1-\delta$ for all $i \in \mathcal{N}$ for an employee type r with $V^{N^r} > \theta$ (i.e., $V^{N^r} = v_{i^*}$ for some $i^* \in N^r$) either $w_{i^*}^{g^r} > V^{N^r} - \epsilon$ or the employee optimally negotiates with employer i^* in any general equilibrium. In words the previous sentence states that an employee with beliefs that each employer entertains counteroffers with probability of at least $1-\delta$ and has a match value with at least one employer that exceeds her outside option will either negotiate with the employer with the highest match value, or must receive an initial offer that yields almost all of the match value from this employer. Recalling that general equilibrium imposes $\hat{\rho}_i^r = \rho_i$ for all $i \in \mathcal{N}$ and all employee types r, showing this statement suffices to prove the desired result, because for small enough ϵ either the employee negotiates with the maximum-match-value employer—in which case she is matched there with probability $\hat{\rho}_i > 1-\delta$—or she does not negotiate with this employer but accepts its offer anyway because its offer is higher than the match value at any other employer: $w_{i^*}^{g^r} \geqslant v_{i^*} - \epsilon > v_j$ for all $j \neq i^*$, where the second inequality holds for small enough ϵ because $v_i \neq v_j$ for all $i \in \mathcal{N}$.

Fix $\epsilon > 0$ and let $v_{i^*} - \epsilon > w_{i^*}^{g^r}$ for all $j \neq i^*$. In the event that $\hat{\rho}_{i^*}^r = 1$, the desired claim follows straightforwardly. In the case that $\hat{\rho}_{i^*}^r \neq 1$, the desired claim follows from Part 2 of Proposition 2, which finds that an employee r optimally negotiates with employer i^* if and only if $\theta_{i^*} \geqslant \frac{w_{i^*}^{g^r} - \hat{\rho}_{i^*}^r v_{i^*}}{1-\hat{\rho}_{i^*}^r}$. By construction, the left-hand side of this inequality is bounded away from $-\infty$ because $\theta_{i^*} \geqslant \underline{\theta} > -\infty$, while the right-hand side of this inequality approaches $-\infty < \underline{\theta}$ as $\hat{\rho}_{i^*}^r \to 1$ because $v_{i^*} > w_{i^*}^{g^r}$. Therefore, for sufficiently small δ, the employee either negotiates with i^* or is initially offered approximately the entire match value at i^* as desired. □

A.1 Employee May Optimally Negotiate any Subset of Offers

As discussed in Section 6.3, negotiation with any subset of employers in an employee's portfolio can be rationalized. The following example demonstrates this, and in doing so shows that finding an optimal portfolio of counteroffers may involve non-trivial considerations, as it is insufficient to "locally" optimize—negotiating with an employer has consequences on the marginal risk courted by the rest of the portfolio.

Example 1. *Let* $N=\{1,2\}$ *and let* $v_1>v_2$*. To rule out uninteresting complexities, let* $w_1<v_1$ *and* $w_2<v_2$*, and let* $\hat{\rho}_1,\hat{\rho}_2\in(0,1)$*. Given Remark 3, there are four potentially optimal portfolios of counteroffers to consider:* $o_1^*=v_1-w_1$ *and* $o_2^*=v_2-w_2$ *(the employee negotiates with both employers),* $o_1^*=v_1-w_1$ *and* $o_2^*=0$ *(the employee negotiates only with employer 1),* $o_1^*=0$ *and* $o_2^*=v_2-w_2$ *(the employee negotiates only with employer 2) and* $o_1^*=o_2^*=0$ *(the employee negotiates either neither employer).*

Option 1: $o_1^*=v_1-w_1$, $o_2^*=v_2-w_2$**.** *Because* $v_1>v_2$*, the employee will match with employer 1 if her offer is not rejected. Therefore, her expected payoff is* $\hat{\rho}_1v_1+(1-\hat{\rho}_1)\hat{\rho}_2v_2$.

Option 2: $o_1^*=v_1-w_1$, $o_2^*=0$**.** *The employee's expected payoff is* $\hat{\rho}_1v_1+(1-\hat{\rho}_1)w_2$.

Option 3: $o_1^*=0$, $o_2^*=v_2-w_2$**.** *By the assumption that the employee will never negotiate with an employer if she has zero probability of matching with that employer, it must be the case that* $v_2>w_1$*. Therefore, the employee's expected payoff is* $\hat{\rho}_2v_2+(1-\hat{\rho}_2)w_1$.

Option 4: $o_1^*=o_2^*=0$**.** *By the assumption that the employee will never negotiate with an employer if she has zero probability of matching with that employer, it must be the case that* $w_1>w_2$*. Otherwise, if* $w_2\geqslant w_1$*, because* $\hat{\rho}_1>0$*, the employee's expected payoff from setting* $o_1^*=o_2^*=0$ *is* w_2 *while the employee's expected payoff from setting* $o_1^*=v_1-w_1$, $o_2^*=0$ *is* $\hat{\rho}_1v_1+(1-\hat{\rho}_1)w_2>w_2$*, where the inequality follows because* $v_1>v_2$ *by the assumption of this example, and* $w_2\leqslant v_2$ *by our ongoing assumption that initial offers are no greater than the match value.*

Because $w_1>w_2$ *and* $o_1^*=0$*, then the employee will optimally negotiate with employer 2 if* $v_2>w_1$*. Therefore, this option is optimal only if* $w_1\geqslant v_2>w_2$*, and the employee's expected payoff is* w_1.

Holding fixed v_1, v_2, w_1, w_2, it is clear that Option 1 is optimal if the risk of negotiating with both employers is sufficiently low. Specifically, there exists $\rho^* < 1$ such that if $\hat{\rho}_1, \hat{\rho}_2 > \rho^*$, the employee will optimally negotiate with both employers.

Similarly, holding fixed v_1, v_2, w_1, w_2 such that $w_1 \geqslant v_2 > w_2$ (recall that these inequalities are necessary for the optimality of option 4), it is clear that Option 4 is optimal if the risk of negotiating with both employers is sufficiently high.

Option 2 is optimal if the risk of negotiating with employer 1 is low and the risk of negotiating with employer 2 is high. Specifically, suppose the employee negotiates with employer 1. Because $v_1 > v_2 \geqslant w_2$, the employee will be matched to employer 1 with probability $\hat{\rho}_1$, i.e., if her counteroffer to employer 1 is accepted. With the complementary probability, her offer is rejected. In the event she is rejected by employer 1, she earns $\hat{\rho}_2 v_2$ in expectation by negotiating with employer 2, and w_2 for sure if she does not negotiate. Therefore, for $\hat{\rho}_1$ sufficiently small, and $w_2 > \hat{\rho}_2 v_2$, option 2 is optimal, i.e., the employee will "risk it" with employer 1 and "play it safe" with employer 2.

Option 3 is similar to option 2, except that employee uses employer 1's offer as a safety, and negotiates instead with employer 2. Recall our previous argument that this option is optimal only if $v_2 > w_1$. Therefore, the employee will be matched to employer 2 with probability $\hat{\rho}_2$, and with the complementary probability, her offer is rejected. In the event she is rejected by employer 2, she earns $\hat{\rho}_1 v_1$ in expectation by negotiating with employer 1, and w_1 for sure if she does not negotiate. It is tempting, mirroring the argument surrounding the optimality of option 2, to reason that option 3 is optimal if $\hat{\rho}_2$ is sufficiently small, $v_2 > w_1$, and $w_1 > \hat{\rho}_1 v_1$.

However, an additional consideration is necessary here because employer 1 is the highest-value employer that the employee will be matched to following a successful negotiation. In other words, the decision of whether or not to negotiate with employer 1 cannot be determined only by conditioning on rejection from employer 2 and "locally" optimizing. To "globally" optimize her portfolio, the employee must consider negotiating with employer 1.[67] *If she does so, she would not negotiate with employer 2 only if $w_2 \geqslant \hat{\rho}_2 v_2$. Therefore, if $w_2 \geqslant \hat{\rho}_2 v_2$, she optimally selects Option 3 if and only if her expected utility from Option 3 exceeds that from Option 2: $\hat{\rho}_2 v_2 + (1 - \hat{\rho}_2) w_1 \geqslant \hat{\rho}_1 v_1 + (1 - \hat{\rho}_1) w_2$.*

[67] Note that the maintained assumptions that $\hat{\rho}_1, \hat{\rho}_2 > 0$, $v_1 - w_1, v_2 - w_2 > 0$ and $v_1 > v_2$, mean that the employee cannot optimally negotiate with neither employer if $v_2 > w_1$; she would be better off negotiating with only employer 1. Because we have previously argued that $v_2 > w_1$ is a necessary condition for the optimality of Option 3, it must be that if Option 3 is potentially optimal, then Option 4 is not.

Similarly, conditional on negotiating with employer 1, she would negotiate with employer 2 if $w_2 < \hat{\rho}_2 v_2$*. Therefore, if* $w_2 < \hat{\rho}_2 v_2$*, she optimally selects Option 3 if and only if her expected utility from Option 3 exceeds that from Option 1:* $\hat{\rho}_2 v_2 + (1-\hat{\rho}_2) w_1 \geqslant \hat{\rho}_1 v_1 + (1-\hat{\rho}_1)\hat{\rho}_2 v_2$*.*

A.2 The Proposed Bargaining Model Rationalizes Experimental Finding on Initial Misperceptions

Evidence from our extensive-margin experimental treatment arm shows that employees who are initially pessimistic about the viability of bargaining experience the largest gains from receiving accurate information (see Figure 4). This short section rationalizes this finding in the context of our model.

Consider a labor market $(\mathcal{N}, \Theta, F, \{v_i\}_{i\in\mathcal{N}})$ in partial equilibrium,[68] and consider two employee types r and r'. We assume that the two employee types are identical in every way, except that they have different beliefs, r has accurate beliefs while r' is pessimistic about the efficacy of bargaining at all employers: $N^r = N^{r'}$, $\theta^r = \theta^{r'}$, $g^r = g^{r'}$, but $\hat{\rho}^r = \rho$ while $\hat{\rho}^{r'} < \rho$.

It follows straightforwardly that employee r's expected payoff (weakly) exceeds that of employee r', as r optimizes her portfolio using accurate information, and r' optimizes using inaccurate information. This observation leads to the following conclusion about how eliminating misperceptions about extensive-margin uncertainty reduces inequality.

Claim 3. *Suppose an information treatment (for example, one similar to that in our experiment) induces correct beliefs about the viability of negotiation in all employees. Then an employee with misperception in her prior belief will benefit relative to an otherwise-identical employee with accurate prior beliefs.*

[68] Recalling Proposition 3, the labor market cannot be in general equilibrium if different employee types have different belief vectors, as this would violate rational expectations.