Female-Targeted Hiring Subsidies, Firm Learning, and Women's Employment.*‡

Mimosa Distefano, Lorenzo Incoronato, Anna Raute

March 2026

Abstract

Women with extended employment interruptions often face difficulties re-entering the labor market. A possible explanation is that employers are uncertain about the productivity of workers with long career gaps. This paper studies whether hiring subsidies can encourage firms to hire from this group. We analyze an Italian policy that temporarily reduced payroll taxes for firms hiring women from non-employment. Combining administrative employer-employee data with an event-study design, we show that following subsidy adoption, firms persistently increase hiring of women with career breaks, including mothers. Additional evidence suggests that the subsidy induces firms to learn about the productivity of workers from the target group, potentially reducing informational frictions in hiring. Women hired under the subsidy also experience stronger long-run labor market attachment, suggesting persistent improvements in labor market integration.

*Distefano: London School of Economics, CEP, RFBerlin, email: m.distefano@lse.ac.uk; Incoronato: University of Naples Federico II, CSEF, CESifo, RFBerlin, email: lorenzo.incoronato@unina.it; Raute: Queen Mary University of London, CEPR, CESifo, RFBerlin, email: a.raute@qmul.ac.uk

†The realization of this article was made possible thanks to the "VisitINPS Scholars" program. This research uses administrative data from the Italian Social Security Institute (INPS). Access to the data is subject to the Institute's confidentiality and renewal procedures. The findings and conclusions expressed are solely those of the authors and do not represent the views of INPS.

‡We are grateful to Paola Onuchic for her generous guidance and helpful discussions, which were instrumental in the development of the conceptual framework. We thank Naoki Aizawa, Thomas Le Barbanchon, Jake Bradley, Antoine Ferey, Giuseppe Ippedico, Paolo Naticchioni, Joao Ramos, Johanna Rickne, Raffaele Saggio, David Seim, Gabriel Ulyssea, Silvia Vannutelli and seminar participants at AIEL, BAEM Meeting on Economic Policy, Bocconi Axa Gender Lab Seminar Series, CEP Labour Market Workshop, CESifo, CReAM Brown Bag seminar and Annual Conference, EEA, EALE, ESWME, IFLAME Seminar, Lisbon School of Economics and Management, Nottingham University, RES and SES Conference, SIEP, Surrey University, University of Milan, USI Lugano, Visitinps Seminar and Annual Conference for helpful comments and suggestions.

1 Introduction

Women remain underrepresented in labor markets around the world. A key driver of persistent gender inequalities is the arrival of children, as it often leads women to take career breaks (Adda, Dustmann, & Stevens, 2017; Angelov, Johansson, & Lindahl, 2016; Kleven, Landais, & Sogaard, 2019). Policy efforts have largely focused on supply-side interventions, such as parental leave and subsidized childcare, which have had limited success in strengthening women's long-run labor force attachment and earnings (Albanesi, Olivetti, & Petrongolo, 2023; Olivetti & Petrongolo, 2008).

In contrast, less attention has been given to demand-side forces, in particular employers' hiring decisions. Employers' reluctance to hire workers with career interruptions, especially when such interruptions are perceived as a signal of lower productivity, may constrain women's labor market opportunities (Arulampalam, 2001; Eriksson & Rooth, 2014; Kroft, Lange, & Notowidigdo, 2013).[1] These concerns may be particularly salient for mothers: sociological research shows that they are perceived as less competent and committed, and consequently face lower hiring probabilities (Correll, Benard, & Paik, 2007; Weisshaar, 2018). If firms are imperfectly informed about the productivity of women returning after career breaks, or if gender stereotypes shape employers' beliefs about their productivity, policies that stimulate demand for this group may help increase women's labor market participation by encouraging firms to experiment with—and learn about—this group.

This paper studies whether a hiring subsidy targeted at women with extended employment interruptions can improve their labor market outcomes by encouraging firms to hire from this group. Female-targeted hiring subsidies have been implemented in several OECD countries to promote women's employment. We study the Italian "Bonus Donne" policy introduced in 2013, which provided a temporary and sizable reduction in employer social security contributions for hiring women from non-employment. Guided by a conceptual framework of employer learning under uncertainty about worker productivity, we study how firms adjust their hiring behavior

[1] Recent evidence suggests that general human capital depreciation during unemployment spells is small (Cohen, Johnston, & Lindner, 2025), implying that perceived productivity—rather than actual skill loss—may play a central role in firms' hesitancy to hire workers with employment interruptions.

after first using the the subsidy and whether initial experiences with subsidized hires affect subsequent hiring decisions. We find that firms using the subsidy increase hiring of women with extended career interruptions, including mothers. Importantly, this did not come at the expense of firm performance. We also provide evidence consistent with firm learning about the productivity of workers in the target group. Finally, we show that women hired under the subsidy experience stronger long-run labor market attachment.

Italy provides a well-suited setting for our analysis because gender differences in the labor market are particularly pronounced: it has one of the lowest female employment rates and the largest gender employment gap in the European Union (Eurostat, 2025). Consequently, increasing female labor force participation may generate substantial economic gains (Carta et al., 2023). The Italian setting also offers a key empirical advantage. We draw on rich matched employer-employee records from the Italian Social Security Institute (INPS), covering the universe of private-sector workers and firms. A unique feature of the data is that we observe whether a worker is hired through the subsidy, which enables us to measure both worker- and firm-level take-up and to track subsidized firms and workers before and after adoption.

We present a conceptual framework linking hiring subsidies to employer learning about worker productivity. The model builds on recent work showing that employers update their beliefs through direct experience with workers (e.g., Benson & Lepage, 2025; Lepage, 2024; Li, Raymond, & Bergman, 2025). In our setting, firms are uncertain about the productivity of workers with extended employment interruptions and may therefore be reluctant to hire them. This uncertainty concerns both the productivity of individual applicants and the match quality of workers from this group more broadly—for instance, women returning after long career gaps. When such uncertainty discourages hiring, firms may fall into a "learning trap" in which workers from the target group are rarely hired, and firms have little opportunity to update their beliefs about their productivity. A temporary hiring subsidy can shift incentives enough to induce experimentation with these workers, generating new information about their productivity. The framework implies that (i) hiring from the target group under the subsidy can lead to persistent changes in firms' hiring behavior, and (ii) these effects depend on the quality of firms' initial

matches with subsidized workers.

To examine both the short- and long-run changes in firms' hiring dynamics following their first subsidized hire under the Bonus Donne, we implement a staggered event study design comparing firms that use the subsidy to firms that hire women in the same year, but not through the program. To account for differences between these two groups, we combine this approach with a matching strategy that pairs treated firms–defined as those that hire at least one worker under the subsidy—with observationally similar control firms. Specifically, we match each treated firm to a comparable control firm based on predetermined characteristics including firm size, average wages, the share of female employees, and the number of hires prior to subsidy adoption.

We find that treated firms increase their hiring of women with lengthy labor market interruptions, including mothers, relative to control firms. The effect is sizable in the year of initial subsidy use—treated firms hire roughly 17% more women with extended employment gaps—and remains positive in subsequent years, with treated firms continuing to hire about 3% more from this group than control firms. We observe a similar increase in hiring for mothers. These patterns are robust across alternative specifications that exploit variation in the timing of subsidy adoption across treated firms. Despite hiring workers with extended labor market interruptions, we find no evidence that firms' labor productivity declines following subsidy adoption.

What explains why firms initially avoided hiring these workers? According to our conceptual framework, firms may be uncertain about the productivity of workers with extended employment interruptions—due to limited past exposure to this group or stereotypes about workers returning after career breaks—and the hiring subsidy can induce firms to experiment and learn about their productivity. If firms learn from interacting with workers, future hiring decisions should depend on the quality of the initial match. We test this implication by examining whether hiring evolves differently across firms that experience better versus worse initial matches with subsidized workers. Consistent with employer learning, firms with better initial matches subsequently hire more women with extended employment interruptions and more mothers.

Finally, we shift the focus to workers hired under the subsidy and investigate their employment

trajectories after hiring and beyond the expiration of the subsidy. Leveraging our ability to follow workers across firms, we implement a matched worker-level event-study design that compares women hired under the subsidy with observationally similar women hired in the same period at matched control firms. We document positive effects on employment probabilities for Bonus Donne workers that persist beyond the subsidy period, indicating that the policy strengthened long-term labor market attachment. Even five years after hiring, subsidized workers are about 8 percentage points more likely to be employed than comparable women at control firms. Importantly, these gains reflect not only higher retention at the hiring firm, but also stronger attachment to the broader labor market, as many treated workers transition into jobs at other firms. We also find no evidence that the Bonus Donne was passed through to workers in the form of higher wages, consistent with Rubolino (2025).

This paper contributes to several strands of the literature. First, it offers a new perspective to the literature on hiring subsidies. Hiring subsidies have long been implemented in many countries to boost employment among low-wage workers (Katz, 1998), or as a countercyclical tool during recessions (Neumark, 2013; Neumark & Grijalva, 2017). The hiring subsidy we analyze is instead targeted at women with extended labor market interruptions, independently of their wage level or the business cycle. While earlier studies relied largely on regional or survey data and focused on aggregate outcomes (e.g., Bishop, 1981; Perloff and Wachter, 1979), recent work has provided micro-level evidence on the effects of hiring subsidies. Cahuc, Carcillo, and Le Barbanchon (2019) document positive employment effects of a French subsidy targeting low-wage workers, whereas Jain, Mommaerts, and Weaver (2025) draw on administrative records from Wisconsin to find that the U.S. Work Opportunity Tax Credit largely operated as a transfer to firms without increasing employment or hiring.[2] In contemporaneous work, Rubolino (2025) focuses on the tax incidence of the Bonus Donne program, rent-sharing within firms, and the policy's fiscal implications, and estimates that most of the program's costs are

[2]There is a sizable literature on the effects of permanent reductions in payroll taxes (e.g, Benzarti and Harju, 2021a, 2021b; Bíró et al., 2024; Huttunen, Pirttilä, and Uusitalo, 2013; Ku, Schönberg, and Schreiner, 2020; Saez, Matsaganis, and Tsakloglou, 2012; Saez, Schoefer, and Seim, 2019, 2021). Unlike a cut to payroll taxation, which covers both new hires and incumbent workers and is typically applied automatically, hiring subsidies apply only to new hires and require active take-up. Therefore, findings from the payroll-tax literature are only partially informative about hiring subsidies. In addition, none of these studies address gender-specific aspects.

ultimately recouped. Our paper focuses instead on how hiring subsidies can lower the cost of experimentation and serve as a *tool* for firms to learn about the productivity of disadvantaged groups, thereby shaping firms' hiring dynamics. This learning channel, highlighted in recent work on employer learning Lepage (2024), has received little attention in the hiring subsidy literature. One exception is Barham, Cadena, and Turner (2023), who study a supported work program in Colorado and provide evidence consistent with employer learning about the match quality of individual workers. In contrast, we study a hiring subsidy targeted at women with extended employment interruptions and examine how firms' initial exposure to subsidized hires from this group affects subsequent hiring decisions for the group more broadly.

Second, our paper offers a complementary perspective to the literature on regulatory policies aimed at increasing labor market representation, such as gender quotas (e.g., Bertrand et al., 2018; Maida & Weber, 2022; Miller, Peck, & Seflek, 2022) and affirmative action policies targeting minorities (e.g., Coate & Loury, 1993; Fryer & Loury, 2013; Miller & Segal, 2012; Miller, 2017).[3] While quotas and affirmative action policies rely on regulatory requirements to promote representation, hiring subsidies instead operate through financial incentives for employers to hire disadvantaged workers. Evidence on affirmative action—largely based on U.S. programs targeting racial minorities—shows that even temporary interventions can have persistent effects on workforce composition (Miller & Segal, 2012; Miller, 2017), although recent evidence suggests that such effects may not generalize to more recent periods or settings (Amano-Patiño, Contractor, & Aramburu, 2025). Our analysis complements this literature by studying a hiring subsidy and highlighting an experiential learning mechanism through which firms' initial exposure to the targeted group can shape subsequent hiring behavior.

Finally, this study relates to the literature on experiential learning in hiring. A growing body of work shows that employers' hiring decisions are shaped by their personal experiences, particularly with underrepresented groups (e.g., Benson & Lepage, 2025; Lepage, 2024; Leung, 2018). Stereotypes or inaccurate beliefs about minority applicants' productivity can impede learning (Bohren, Imas, & Rosenberg, 2019), and negative experiences with minority hires may discourage

[3] On the theoretical rationale for such policies, see Fryer and Loury (2013), who show that valuing diversity can enhance social welfare even when it entails short-run efficiency costs.

future hiring from those groups, thereby preventing the correction of biased beliefs (Benson & Lepage, 2025).[4] Studying the role of hiring algorithms, Li, Raymond, and Bergman (2025) show that exploration-based screening algorithms encourage employers to learn about workers with less common characteristics and thereby improve both the diversity and quality of hires. This insight relates closely to the mechanism we propose: hiring subsidies reduce the short-run cost of experimentation, enabling firms to learn about the potential productivity of disadvantaged workers.

The paper is organized as follows. Section 2 describes the institutional setting and the policy. Section 3 presents the data and provides descriptive statistics. Section 4 introduces our theoretical framework. Section 5 reports the empirical strategy, the main estimates and robustness checks, while Section 6 presents evidence on firm learning. Section 7 explores worker-level effects. The last Section concludes.

2 Institutional Background

2.1 Gender Gaps in the Italian Labor Market

There has long been a substantial gender divide in Italy's labor market. Women's employment and labor force participation are low compared to other advanced economies: Italy has the lowest female employment rate among EU countries, at 57%, and the largest gender employment gap (Eurostat, 2025). A large share of this gap emerges around childbirth, with 60% of the gender employment gap explained by the child penalty (Kleven, Landais, & Leite-Mariante, 2024). The employment drop following childbirth also implies substantial and persistent earning losses for mothers (Casarico & Lattanzio, 2023). The historically strong employment protection for incumbent jobs in Italy—partly shaped by traditional male breadwinner norms—has been shown to contribute to lower female employment rates (Giavazzi, Serafinelli, & Schiantarelli, 2013).[5] Like in other European countries, the Italian labor market exhibits a dual structure, where

[4] Relatedly, Cullen, Dobbie, and Hoffman (2022) experimentally test interventions providing employers with information about worker performance or reducing perceived hiring risks for workers with criminal records, and find that these increase employers' willingness to hire from this group.

[5] Gender gaps in labor market outcomes in Italy are further exacerbated by less egalitarian gender norms, especially in Southern regions (Campa, Casarico, & Profeta, 2010).

incumbent workers on open-ended contracts ("insiders") enjoy a high degree of employment security, while new hires ("outsiders") are often offered temporary contracts with limited unemployment protection (Boeri, 2011). This leads to labor market segmentation, with women disproportionately over-represented in outsider categories and experiencing more intermittent labor market participation due to employment in more precarious contracts (Carta et al., 2023).

2.2 The "Bonus Donne"

At the height of the European sovereign debt crisis, many European countries, including Italy, implemented structural reforms aimed at stimulating employment and economic growth. On June 28, 2012, the Italian government passed a comprehensive reform package (Law n. 92/2012 *"Provisions on labor market reform from a growth perspective"*). This legislation introduced various measures intended to create an inclusive and dynamic labor market, increase employment, and improve job stability and job quality, with the aim of reducing unemployment in the long run..[6]

The hiring subsidy analyzed in this paper was one of the measures introduced as part of this reform package. It came into force on January 1, 2013 and was dubbed "Bonus Donne" ("Bonus Women") by the media. The low female employment rate was an important policy priority for the government at the time: *"Ensuring the full inclusion of women in all areas of working life is a matter that cannot be postponed. It is necessary to address issues concerning the reconciliation of work and family life [...] as well as to consider the introduction of preferential taxation for women"* (Interview to Prime Minister Monti, La Repubblica, 2011). The subsidy granted a 50% temporary reduction in employers' social security contributions for hiring non-employed women. Considering that employer-borne social security contributions account for about 20% of the wage in Italy, the Bonus Donne provided a roughly 10% cut to employers' labor costs of new hires, comparable in size to hiring subsidies in the French and U.S. contexts (Cahuc, Carcillo, & Le Barbanchon, 2019; Neumark & Grijalva, 2017).[7]

[6] Appendix A.1 provides additional detail on the policy package and the political context.

[7] Both employers and employees are subject to payroll taxes in Italy. Payroll tax rates have been stable since 2005 at about 20.6% for employers and about 10% for employees.

The subsidy varied along several margins, including contract type, eligibility duration, and worker characteristics. Firms receive the subsidy for 12 months when they hire a worker on a fixed-term contract, and for 18 months when they hire a worker on an open-ended contract or convert a fixed-term contract into an open-ended one within 12 months. The extension of the benefit period for open-ended contracts reflects the policymaker's explicit aim of promoting stable employment and labor market attachment among non-employed women. The subsidy applied to private-sector workers on both full-time and part-time contracts, but excluded apprenticeships and managerial positions. The total Euro amount of the subsidy per worker was uncapped until 2022, which lies outside our analysis period. The eligible target group consisted of individuals who had been out of employment for at least 24 months, regardless of whether they were registered as unemployed. The Bonus Donne also applied to new labor market entrants. The 24-month non-employment requirement was more lenient for some categories of workers.[8] [9]

The legislation also imposed several requirements on firms, entailing non-trivial administrative costs. First, firms were required to submit a separate online claim to the Italian Social Security Institute for each subsidized hire, and employers' social security contributions were reimbursed ex post at the start of the following tax year. Second, the subsidy could be applied only if no other worker with a legally granted preferential hiring right was available for the position.[10] Finally, the subsidy was available only to firms that had not experienced a net decrease in employment—excluding worker resignations and retirements—during the 12 months prior to claiming the Bonus Donne. Net job growth requirements are not uncommon in hiring subsidies—for instance, the U.S. New Jobs Tax Credit was an "incremental employment" subsidy (Katz, 1998)—but they often impose significant administrative burdens on firms and limit the

[8] The non-employment period was reduced to 6 months for female workers hired in male-dominated occupations or sectors, or for residents in municipalities eligible for EU Structural Funds. Male-dominated occupations and sectors are identified each year by ministerial decrees as occupations and sectors where the gender employment gap is at least 25% larger than the average. There is little time variation in these lists. EU Structural Funds municipalities were first identified in 2008, and updated in 2014. Most of these municipalities are located in Southern Italy.

[9] For women over the age of fifty, subsidy eligibility began after 12 months of registered unemployment. Only for this target group, the law requires a formal unemployment status. In this case, the subsidy also applies to men over fifty. However, few men were hired through the subsidy, and these hires were concentrated in specific, male-dominated industries such as construction and waste management. Hiring-subsidized men are excluded from the analysis.

[10] In Italy, a worker previously employed under a fixed-term contract for more than six months has a preferential right to be rehired by the same firm within six months of the contract's end.

pool of eligible users (Neumark, 2013).

These provisions, coupled with the fact that the subsidy targeted non-employed women—a group that some firms may be reluctant to hire (Hamersma, 2003; Katz, 1998)—help explain the relatively low take-up of the Bonus Donne. Using the data described in Section 3, we calculate that 3.8% of firms used the Bonus Donne to hire at least one woman between 2013 and 2019. The share of non-employed women hired under the subsidy was initially quite low—possibly because few firms were aware of the Bonus Donne's introduction—then gradually increased over time, reaching 8% in 2019 (Appendix Figure A.1).[11] Reflecting growing interest from firms, the Italian government has retained the Bonus Donne and strengthened its generosity over time. Since 2022, the subsidy has been increased to cover 100% of employers' social security contributions, although subject to a cap.

3 Data and Descriptive Evidence

This section describes the data used in the analysis and presents descriptive statistics on the workers and firms that received subsidies under the Bonus Donne.

3.1 Data Description and Sample Selection

We use administrative matched employer-employee data from the Italian Social Security Institute (INPS), which we access through the VisitINPS Scholars program. The data cover the universe of private-sector workers and firms subject to the Italian social security system, thus excluding the self-employed and public-sector employees. Several features of this dataset make it uniquely suited to our research questions.

Take-up. Crucially, and unlike in most other contexts, we observe subsidy take-up at both the individual and firm levels, as the data directly report whether a worker was hired under the Bonus Donne. This enables us to characterize hiring-subsidized workers and firms relative to the

[11]The Bonus Donne was introduced alongside more controversial and prominent policies that dominated public discourse (see Appendix A.1). Previous studies have also documented limited awareness of hiring subsidies—for example, only 34% of U.S. firms were aware of the U.S. New Jobs Tax Credit (Perloff & Wachter, 1979).

average Italian worker and firm. To the best of our knowledge, no other countrywide dataset used in the study of hiring subsidies allows simultaneous observation of both worker- and firm-level take-up, along with detailed information on the firm's workforce.[12]

Workers. We observe workers' employment histories since labor market entry, along with their labor earnings and the number of weeks worked each year (in full-time equivalent), which allows us to compute a worker's full-time-equivalent weekly wage.[13] The data also report information on the employment contract (part-time versus full-time, fixed-term versus open-ended), broad occupational category (apprenticeship, blue-collar, white-collar, and managers), and demographic characteristics such as age, gender, place and year of birth, year of labor market entry, and municipality of work.[14]

Defining mothers. Maternity leave spells are available in the INPS data starting in 2005, which enables us to identify births among women employed prior to childbirth, as is common in studies using administrative employer-employee data (e.g., Lalive et al., 2014; Schönberg & Ludsteck, 2014). Throughout our period of analysis, paid mandatory maternity leave in Italy lasts five months—remunerated at 80% of the salary—of which up to two can be taken before childbirth. Pregnant workers also benefit from employment protection. Voluntarily exiting the labor force without taking leave results in the loss of both job protection and maternity benefit entitlements. Consequently, nearly all women employed prior to childbirth take maternity leave.[15] We define mothers as women who experienced a maternity leave event in the previous four years, restricting attention to mothers with young children—a key group of interest for a policy targeting women's re-entry into the labor market.

[12]For example, Cahuc, Carcillo, and Le Barbanchon (2019) use firm-level administrative data for France, but lack worker-level information. A recent exception is Jain, Mommaerts, and Weaver (2025), which utilizes matched data for Wisconsin.

[13]The "full-time equivalent weeks of work" variable exploits information on the number of hours worked in a month (available to INPS, but not to researchers).

[14]Since 2010, more detailed information on occupations (based on the ISCO classification) is available for a subset of workers observed upon contract changes (e.g., layoffs, promotions). Due to the high incidence of missing values for this detailed occupation variable—especially during the early reform years—we do not use it in our analysis.

[15]Paid paternal leave in Italy is much shorter—a maximum of two to five working days during our analysis period. Parents could take additional leave, up to ten months, until their child is 12 years old, with 30% replacement rate.

Firms. Worker-level data can be matched to firms using unique firm identifiers. The firm database reports additional information such as sector (up to 6-digit), location, and firm age. We also obtained balance sheet data from the Cerved Group databases, recording information on total costs, revenues, assets, value added, and profits. The Cerved data are available until 2018 for limited liability companies only and can be merged into the INPS data using a unique national tax firm identifier (*codice fiscale*).

Sample construction. We build a panel dataset consisting of one observation per worker per year between 2008 and 2019—approximately 160 million observations—which forms our worker sample and which we use to construct workforce and hiring composition at the firm level.[16] We restrict the analysis to workers aged 18 to 67 and drop sectors that are only partially represented in the INPS data such as agriculture, electricity and water supply, and international organizations. We drop the (very few) cases where earnings are negative or missing, and where the gender variable is missing. Finally, we exclude one-employee firms. After data cleaning and incorporating information on take-up, we calculate that about 180,000 women have been hired through the Bonus Donne between 2013 (the year the subsidy was introduced) and 2019 (our last year of observation), and around 26,500 firms hired at least one female worker under the subsidy over the same period.[17]

3.2 Descriptive Statistics

We now turn to the focus of our analysis—the characteristics of firms and workers covered by the Bonus Donne. Table 1 contrasts average characteristics of hiring-subsidized firms (Bonus Donne firms, Column 1)—i.e., those that used the Bonus Donne at least once between 2013 and 2019—with those of the average Italian private-sector firm (Column 2). Each variable is measured in 2012, the year before the introduction of the policy. Bonus Donne firms differ from the average Italian firm in multiple aspects. They are larger (48 vs. 18 employees on average;

[16] If a worker holds multiple jobs in a year, we select the job associated with the highest number of weeks worked; if these are equal, we keep the job with the highest weekly wage. Because maternity leave data are available from 2005 onward, the analysis will begin in 2008; maternity histories are observed over a slightly shorter window in the first year, and over the full four-year window thereafter.

[17] We conclude our analysis in 2019, as Italy was severely impacted by the Covid-19 pandemic, which led to a nationwide lockdown in early 2020.

see also Table A.1 Panel (a) for details on the firm size distribution) and employ a higher share of women (57%) compared to the average firm in the data (44%)—suggesting that these may be more female-friendly firms. Indeed, Bonus Donne firms have a have a lower share of full-time contracts (63% versus 70% in other firms), consistent with greater use of part-time employment, and a higher proportion of female workers among their new hires (54% versus 42% in other firms). In addition, Bonus Donne firms hired more workers between 2010 and 2012 (both in absolute numbers and relative to firm size), suggesting differential growth trajectories. They, however, pay about 3% lower wages, both on average as well as for female workers, and have lower average productivity (measured as value-added per worker). Finally, Bonus Donne firms are more likely to be located in the South of Italy—see Appendix Table A.1 Panel (b)—and to operate in services such as wholesale and retail trade, hospitality, and administrative services (Appendix Table A.2).

We then turn to subsidized female workers and compare them with the average female hire between 2013 and 2019 (Columns 1 and 2 in Appendix Table A.3, respectively). Bonus Donne hires are, on average, older and earn slightly higher wages in both their current and previous jobs (if any) than the average female hire. While these differences may hint at positive selection, Bonus Donne workers have, on the other hand, substantially longer employment interruptions, are slightly more likely to be mothers of young children (despite being older on average), and are more likely to be hired under a fixed-term contract—indicating weaker prior labor market attachment than the average female hire. Several of these characteristics—most notably the length of prior non-employment spells and motherhood status—feature prominently in our later analysis of how firms adjust their hiring behavior following subsidy adoption.

4 Model

In this section, we outline a conceptual framework that clarifies how a one-time hiring subsidy may affect firm hiring decisions persistently, building on recent work on experience-based learning (Benson & Lepage, 2025; Lepage, 2024). Appendix A.4 provides more detail. Employers are uncertain about the firm-specific productivity of individual workers and take a worker's

group membership into account when making hiring decisions; in addition, we assume that they are uncertain about the productivity of different groups at the firm. This uncertainty may be particularly pronounced for disadvantaged groups such as women with long non-employment spells—those targeted by the Bonus Donne—who are often perceived as "outsiders" to the labor market and for whom firms may have limited information (e.g., Benson, Board, & Vehn, 2024; Cornell & Welch, 1996; Lundberg & Startz, 1983). Through the model, we show that a hiring subsidy can nudge firms to experiment by hiring from the targeted group and that such experimentation can generate learning about the group's productivity at the firm, potentially increasing hiring from that group in the long run.

We assume that a worker i from group g working in firm j has productivity

$$y_{igj} = \bar{y}_{gj} + \theta_{gj} + \theta_i,$$

where $\bar{y}_{gj}$ represents a known baseline fit of group g to firm j, θ_{gj} is an unknown group-firm component of productivity (i.e., match quality), and θ_i represents an unknown, idiosyncratic component of individual i. We assume that $\theta_{gj} \sim \mathcal{N}(\mu_{gj}, 1)$ and $\theta_i \sim \mathcal{N}(0, 1)$. Thus, in contrast to models of statistical discrimination (e.g., Aigner & Cain, 1977; Arrow, 1973; Phelps, 1972), where employers act on true group-level differentials, we assume that the firm is uncertain about *both* the individual's productivity θ_i and the group-firm component of productivity θ_{gj}. These assumptions distinguish our framework from standard models of statistical discrimination by allowing firms not only to learn about individual productivity, but also about the groupâfirm component of productivity.

The model spans two periods. In each period, firms may hire and thereby learn about the targeted group.

The hiring process. The hiring process takes place in two stages. First, the firm decides whether to interview a candidate. The firm observes only the worker's group membership g and decides whether to interview based on its prior belief about the expected productivity of group g at firm j (i.e., $\bar{y}_{gj}$ and μ_{gj}). Since interviews are costly (Kuhn & Yu, 2021), the firm

interviews a candidate if the expected value of conducting the interview for workers from that group exceeds the interview cost, $V_g \geqslant \kappa > 0$. We define V_g below.

If the firm conducts the interview, it observes a noisy signal of worker i's productivity,

$$\tilde{y} = y_{igj} + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. After observing $\tilde{y}$, the firm updates its belief about the worker's productivity, forming a posterior expectation $\hat{y}_{igj} = \mathbb{E}[y_{igj} \mid \tilde{y}]$.[18] We assume that the worker's wage w is proportional to their expected productivity at the firm, $w = \alpha \hat{y}_{igj}$, for some fixed $\alpha \in (0, 1)$. The firm's surplus from hiring the worker is $s = (1 - \alpha)\, \hat{y}_{igj}$, and the firm hires the worker if this surplus is positive.

The ex-ante expected value of interviewing a candidate from group g is

$$V_g = \mathbb{E}_g \left[\max \left\{ 0, (1 - \alpha) \hat{y}_{igj} \right\} \right],$$

where the expectation is taken with respect to the firm's beliefs about productivity in group g, given the firm's current information about group-firm fit ($\bar{y}_{gj}$ and μ_{gj}) and anticipated signal noise (σ^2). We assume that, in the absence of the subsidy, the firm's prior belief about the targeted group's productivity is sufficiently low that $V_g < \kappa$, creating a "learning trap" in which no interview—and hence no updating and no hiring—occurs.

Now consider a hiring subsidy that covers a proportion $\beta \in (0, 1)$ of the wage. This lowers the firm's effective wage cost to $w = (1 - \beta)\alpha \hat{y}_{igj}$, raising the firm's surplus from hiring the worker to $s = [(1 - \alpha) + \beta\alpha]\, \hat{y}_{igj}$. Accordingly, the ex-ante expected value of interviewing a worker from group g increases to

$$V_g = \mathbb{E}_g \left[\max \left\{ 0, ((1 - \alpha) + \beta\alpha)\, \hat{y}_{igj} \right\} \right].$$

By raising V_g, the subsidy increases the expected return from interviewing candidates from the

[18] See Appendix A.4 for the full expression of the firms' updated beliefs, following standard Bayesian updating rules.

targeted group, encouraging firms to experiment by interviewing and hiring some workers from this group.[19]

Learning. Through the signal observed during the interview, the firm learns not only about the productivity of the individual worker, but also about the group-firm component of productivity, θ_{gj}. This learning is especially important in our context, as it connects the firm's initial interview—and any subsequent hiring decisions induced by the subsidy—to the firm's subsequent behavior toward the targeted group. The posterior expectation of θ_{gj} given the interview signal is:

$$\hat{\mu}_{gj} = \mathbb{E}\left[\theta_{gj}|\tilde{y}\right] = \frac{1}{2+\sigma^2}(\tilde{y} - \bar{y}_{gj}) + \frac{1+\sigma^2}{2+\sigma^2}\mu_{gj}. \tag{1}$$

The firm's belief about the group-firm component of productivity θ_{gj} increases when the "surprise" term $(\tilde{y} - \bar{y}_{gj})$ exceeds its prior mean μ_{gj}, and decreases when it is smaller. This learning process implies that a one-time subsidy can generate lasting effects. By nudging firms to consider—and ultimately hire— workers from a group they would otherwise avoid, the subsidy generates new information about group-firm match quality, potentially breaking the "learning trap" and leading firms to update their beliefs about this match quality. The extent of this updating depends on the realized signal: firms with better-than-expected initial experiences revise their beliefs more positively.

Second period. At the start of the second period, firms differ in what they have learned about their match quality with group g, denoted θ_{gj}. Some firms may have experienced *no learning* if they chose not to interview in the first period. For these firms, second-period hiring mirrors that of the first period, except for any policy changes (e.g., adjustments in the hiring subsidy) that alter the expected value of interviewing V_g.

In contrast, firms that interviewed a candidate in the first period observed a signal $\tilde{y}$ and updated their beliefs about θ_{gj} based on the "surprise" term $(\tilde{y} - \bar{y}_{gj})$. This learning shifts their posterior distribution of θ_{gj} to $\theta_{gj} \sim \mathcal{N}(\hat{\mu}_{gj}, \hat{\sigma}^2)$, where $\hat{\mu}_{gj}$ is given by Equation (1) and $\hat{\sigma}^2 < 1$. These

[19] Importantly, the firm's hiring rule—hiring all candidates with non-negative surplus—remains unchanged. The subsidy raises the expected return to interviewing, thereby shifting firm behavior at the interview margin.

updated beliefs affect second-period interview and hiring decisions. In particular, firms that received more favorable signals (i.e., higher realizations of $\tilde{y}$) in the first period are more likely to interview and hire workers from group g in the second period relative to firms that observed less favorable signals.

To interpret hiring wages in the data, we assume that wages are set at the time of hiring, after the firm observes the signal about the worker's productivity. Under this assumption, the hiring wage—particularly deviations from predicted wages (wage residuals)—reflects the information revealed by this signal and thus proxies for the firm's perceived match quality at the time of hire. This provides the conceptual foundation for our empirical interpretation of wage residuals, as discussed in Section 6.

Implications. Our framework highlights two mechanisms through which a hiring subsidy can affect firm hiring behavior.

1. The hiring subsidy can have persistent effects on firms' decisions to interview—and subsequently hire—from the targeted group if it helps resolve a "learning trap". By nudging firms to experiment with workers from the targeted group, the subsidy facilitates learning about the group-firm match quality.[20]

2. The long-term effect of the subsidy on firm hiring depends on the quality of initial matches: firms that experience higher-quality matches update their beliefs more positively and are more likely to continue hiring from the targeted group than firms whose initial matches are less favorable.

These mechanisms generate key empirical predictions that we test in the data. First, firms exposed to the subsidy should increase hiring from the targeted group. Second, we expect heterogeneous responses across firms depending on the quality of their initial matches with workers from the targeted group. In practice, very poor matches may not lead to hiring and are therefore not observed. Accordingly, since we only observe hired workers, our empirical

[20] In our model, learning can generate persistent effects even if the subsidy is only temporary, consistent with empirical findings by Miller (2017) on temporary affirmative action policies.

comparison is between firms that experienced relatively stronger versus weaker matches among hired workers.

5 Firm Dynamic Responses to Subsidy Take-Up

As documented in Section 3, women hired under the Bonus Donne are characterized by long employment interruptions and differ systematically from the average female hire along several other dimensions. We now examine whether firms' exposure to such workers leads to persistent changes in their hiring behavior. Do firms continue to hire women with extended employment interruptions after their initial subsidized hires, or is such exposure short-lived, with firms reverting to their previous hiring patterns?

5.1 Empirical approach

Answering these questions requires constructing a counterfactual scenario: how would hiring have evolved had the firm not used the subsidy? As described in Section 3.2, Bonus Donne firms differ from the average Italian firm: they tend to be larger, more female-friendly, and possibly on a steeper growth trajectory. These differences may correlate with subsequent hiring patterns and lead to spurious results. To construct an appropriate control group and approximate how hiring would have evolved in the absence of the subsidy, we combine a matching strategy with a staggered event-study design, comparing the evolution of outcomes between Bonus Donne ("treated") firms—where treatment is assigned from the year of the first subsidized hire onward—and observationally similar "control" firms.[21]

Firm matching. We narrow our focus to firms that are observed every year between 2008 and 2019.[22] We define subsidy adoption as the year of the first subsidised hire, and define treated

[21] Our empirical strategy directly exploits take-up information and compares firms that use the subsidy to those that do not, rather than relying on the eligibility margins introduced by the policy (Footnotes 8-9). These margins would exploit variation across, for example, broad (1-digit) sectors, or between areas subject to EU structural interventions and those that are not, before and after the introduction of the Bonus Donne in 2013. However, such variation induces little change in subsidy take-up—particularly in the early years, when overall take-up was low (Section 3)—and may raise additional endogeneity concerns. For these reasons, we focus on comparisons between treated and untreated firms, accounting for differences in key firm characteristics through a matching strategy.

[22] By conditioning on firms operating over the entire period, we ensure that results do not reflect differential firm survival rates. In additional checks, we find no significant differences in survival rates between treated and matched

firms as those that hire at least one female worker through the Bonus Donne between 2013 and 2019. We match each treated firm to a twin control firm using propensity score matching based on an extensive set of pre-determined variables. Specifically, we fit a logit model that relates the probability of subsidy take-up to the following firm-level covariates in the three years before a treated firm's first subsidized hire: average (log) weekly wage, (log) total employment, female share of employment, and quartiles of the number of workers hired.[23] Additionally, we require that control firms hire at least one female worker in the same calendar year in which their matched treated firm hires a worker through the Bonus Donne for the first time. These female hires at control firms serve to characterize the average female hire in the absence of subsidy use.[24]

The resulting sample consists of 38,270 firms, half of which are treated (19,135—corresponding to a matching rate of approximately 72‰).[25] Table 2 shows the balancing properties of the matched sample. Treated and control firms are well matched on average wages, workforce size, and the share of female employees. Additional indicators of workforce composition that were not explicitly targeted in the matching procedure, such as the share of full-time or open-ended contracts, are also quite similar between the two groups, suggesting that the matching procedure yields well-balanced comparison firms.

Firm event study design. We perform the following staggered event-study analysis on the matched sample of firms:

$$y_{jt} = \xi_j + \rho_t + \sum_{k=-4}^{5} \gamma_k \cdot \mathbb{1}[t = t_j^* + k] + \sum_{k=-4}^{5} \beta_k \cdot \mathbb{1}[t = t_j^* + k] \cdot T_j + e_{jt}, \tag{2}$$

where y_{jt} denotes the outcome variable for firm j in calendar year t, T_j is a binary indicator

control firms.

[23]The matching is performed without replacement and with a 0.05 caliper for the propensity score.

[24]We restrict control firms to those that hire at least one female worker in the initial period to ensure a meaningful counterfactual (e.g., when examining mothers, a male control hire would never be a valid comparison) and to improve overall comparability between treated and control firms. Our results are robust—if anything, larger in magnitude—when we relax this restriction and consider control firms that hired workers of any gender in the initial period.

[25]The cohort distribution for treated firms (hence for controls, too) is as follows: 2,498 in 2013, 5,372 in 2014, 1,816 in 2015, 3,170 in 2016, 8,524 in 2017, 8,710 in 2018, and 8,180 in 2019.

taking value one for treated firms and zero for control firms, and $t_j^* \in [2013, \ldots, 2019]$ denotes the year in which treated firm j first hires a worker through the Bonus Donne (for a control firm, this corresponds to the year of first subsidized hire of their matched treated firm). The specification controls for firm-specific unobserved characteristics (ξ_j), aggregate shocks (ρ_t), and cohort-specific effects (captured by the γ_k coefficients). The coefficients of interest are the β_k's, which capture the difference in outcomes between treated and control firms k years before or after the first subsidized hire. These effects are measured relative to the same difference in period $k = -1$, which is normalized to zero. Standard errors are clustered at the firm level.

By comparing outcomes for firms not only in the same calendar year but also in the same period relative to subsidy adoption—aligning control firms to the adoption year of their matched treated firm—we mitigate potential challenges inherent in event-study models with staggered treatment timing (Sun & Abraham, 2021).[26] When reporting our estimates, we refer to the coefficient β_0 as the immediate (enactment) change in hiring in the year of subsidy adoption, while coefficients for subsequent years capture longer-run responses in firms' hiring patterns following adoption.

Threats to identification. The main identifying assumption is that hiring at treated and control firms—matched on key indicators of workforce composition and prior hiring—would have followed parallel trends in the absence of subsidy adoption. To assess the plausibility of this assumption, we examine the evolution of hiring outcomes in the years preceding adoption, as captured by the pre-event coefficients $\beta_{k<0}$. To corroborate our results further, we perform additional exercises that augment Equation 2 with province-by-year or (2-digit) industry-by-year effects—which account for industry-specific and local labor market shocks that could influence both subsidy uptake and hiring. In the most stringent specification, we introduce pair-by-year fixed effects that allow for pair-specific time shocks (where each pair consists of a treated firm and its matched control). As we will show, the estimates remain robust to these additional controls.

[26] Without controlling for period-from-event effects, Equation 2 would implicitly assume homogeneous treatment effects across early and late adopters, potentially distorting the estimates through negative weights. By incorporating a matched control group, we are able to estimate period effects and avoid these distortions—see also Baker, Larcker, and Wang (2022) and Cengiz et al. (2019).

A remaining concern is that unobserved factors may jointly influence subsidy adoption and firms' hiring dynamics in ways that are not fully captured by our extensive set of covariates. To further mitigate concerns about selection into treatment, we conduct a final robustness check that excludes firms that never adopt the Bonus Donne and focuses solely on *treated* firms, estimating staggered event-study models that exploit variation in treatment timing among these firms only. To address treatment effect heterogeneity and the resulting bias, we follow Sun and Abraham (2021) and use the last treatment cohort (2019) as the control group. Our results are robust to this alternative strategy, which strengthens our confidence that the observed changes in firms' hiring behavior following subsidy adoption are not driven by unobserved selection into treatment.

While our empirical specifications and robustness tests yield consistent results, we acknowledge that this approach may still be subject to potential threats, for instance, time-varying unobserved firm shocks that drive subsidy adoption. However, our key goal with this analysis is not to causally identify the impact of the hiring subsidy *per se*, but rather to understand how a subsidy-induced change in the composition of firm hires shapes subsequent firm behavior. In this context, we interpret this firm-level evidence as consistent with the mechanism outlined in the theoretical section, whereby hiring subsidies induce firms to experiment with—and learn about—the targeted group.

Main outcome variables. Our analysis focuses on several outcomes derived from the data described in Section 3. To study the direct effects of subsidy take-up on firm hiring and labor costs, we examine two outcomes: the share of female new hires hired through the Bonus Donne, defined as subsidized female hires divided by total female hires in the year, and the average (log) gross weekly wage of female new hires, measured inclusive of employer social security contributions and thus reflecting the firm's labor cost per hire. To further characterize female new hires, we consider three dimensions: the average number of previous non-employment years of female new hires at the firm; the (log) number of female new hires from long-term non-employment (i.e., a non-employment period of at least three years); and the (log) number of mothers among female hires (see Section 3 for a definition of mothers). Finally, we assess broader firm-level outcomes, including the (log) number of male new hires; (log) total firm employment;

(log) firm value added; and (log) firm value added per (full-time equivalent) worker.[27]

5.2 Results

We begin by presenting the evolution of subsidized female hires in treated firms in Figure 1 Panel (a).[28] In the first year of subsidy use (the event year), 70% of new female hires in treated firms were hired through the subsidy. While subsidy use tends to be concentrated in the initial year, treated firms continue to use the subsidy to a lesser extent in subsequent years, with an average of 6% of new female hires subsidized.

This surge in subsidized hiring at treated firms is associated with a reduction in the average labor costs of new female hires. Figure 1 Panel (b) shows event-study coefficients and corresponding confidence intervals for the average labor cost of new female employees. Prior to the first use of the subsidy, female hires' labor costs at treated and control firms were on parallel trends. In the event year, labor costs for female new hires at treated firms decline by 2.6% compared to control firms.[29] This gap in labor costs appears to persist over time, though estimates become noisier and smaller in magnitude.

We next shed light on the characteristics of these new female hires compared to the average female hire, both in the event year and in the long run. Figure 2 Panel (a) investigates the length of the prior non-employment spell of female new hires at treated and control firms, plotting event study coefficients for the average duration (in years) of non-employment spells. While non-employment spell length of female new hires evolved similarly at treated and control firms in the pre-treatment years, we find a notable difference at the time of treatment. The average non-employment spell for female new hires in the event year was 1.5 years longer in treated firms than in control firms, representing a nearly 50% increase. Crucially, this effect persists

[27] For female hires from long-term non-employment, mothers among female hires, and male hires, the log-outcomes are computed as $\log(y+1)$ to capture both extensive and intensive margin adjustments.

[28] As control firms never use the subsidy, the graph simply plots the share of female workers hired through the subsidy as a percentage of new female hires for treated firms.

[29] Here, we do not control for differential selection among new hires. In Section 7, we examine worker-level outcomes while explicitly accounting for selection (on observables) and show that there is no evidence of wage pass-through: the temporary cut in employer social security contributions was fully retained by firms and not passed on to workers as higher wages. This lack of pass-through aligns with the results of Rubolino (2025) and suggests that the subsidy effectively reduced the cost of labor for firms.

beyond the event year, with treated firms continuing to hire women with longer labor market interruptions in subsequent periods.

Panels (b) and (c) further explore hiring patterns for two groups that may be facing the most significant barriers to re-enter the labor market—long-term non-employed women and mothers. For both outcomes, the pre-event coefficients confirm parallel trends between treated and control firms prior to the first subsidized hire. Treated firms increase their hiring of women with long non-employment spells (of at least three years) in the event year by more than 15%. Again, these effects persist—albeit to a lower extent—in the years following the first subsidized hire, with coefficient of around 3%. Panel (c) shows that treated firms also increase the number of recent mothers among new hires upon first use of the subsidy. The number of newly hired mothers rises by 4% in the event year at treated firms, followed by positive and significant coefficients of around 2% in the subsequent periods.

Taken together, the evidence from Figures 1 and 2 suggests that the Bonus Donne has led firms to broaden recruitment to the targeted group—women with extended employment interruptions. Viewed through the lens of our model, the subsidy increases the expected return to hiring from the group, making it more attractive for firms to hire workers they might otherwise have avoided, who may have been perceived as a poorer match. These findings illustrate how policies designed to improve access to employment can influence firms' hiring patterns by inducing firms to experiment with—and learn about—the targeted group.

We complement the above evidence by examining additional firm-level outcomes, as reported in Figure 3. Panel (a) shows that the hiring of men evolved similarly at treated and control firms prior to the first subsidized hire. In the two to three periods following the first subsidized hire, treated firms hired about 2% *more* men than control firms; the effect then reverts to zero four years later. Consistent with this increase in hiring, Bonus Donne firms also grow more rapidly following the first subsidized hire, ultimately reaching a long-run increase in firm size of about 6%, as shown in Panel (b). This steeper post-event growth path may partly reflect lower labor costs for new hires, which allow firms to expand. Finally, as reported in Panel (c), treated firms experience a long-run increase in value added of around 5%, while productivity per worker

remains unchanged (Panel (d)). Importantly, the absence of any decline in labor productivity suggests that workers hired through the subsidy did not adversely affect firm performance and were therefore no less productive than the average hire, despite having longer labor market interruptions.

5.3 Robustness checks

We now present robustness checks, focusing on a subset of outcomes (average non-employment spell, new hires from long-term non-employment, mother new hires, and firm size).[30] Figure A.3 plots event-study coefficients and corresponding confidence intervals estimated according to Equation 2 after adding i) province-by-year fixed effects, ii) 2-digit industry-by-year fixed effects, and iii) firm match-pair-by-year fixed effects. The first two specifications exploit variation in subsidy use within provinces or industries over time, while the third specification controls for time shocks specific to each matched firm pair, effectively exploiting within-pair variation over time. Event study coefficients are remarkably similar across specifications and to the baseline estimates, suggesting that our results are not driven by differential trends at the province, industry or matched-pair level.

Figure A.4 shows event-study coefficients estimated on a sample restricted to treated firms, using the last treatment cohort (2019) as the control group (Sun & Abraham, 2021). By comparing outcomes for firms that first use the Bonus Donne at different points in time, this analysis addresses potential selection concerns that treated firms may differ from control firms in ways not captured by our matching approach. While point estimates differ in magnitude from the baseline matched-sample estimates—and are in fact larger for most outcomes, albeit with some evidence of pre-trends—our key findings are confirmed in the treated-only sample. This strengthens our confidence that the observed changes in firms' hiring behavior are not primarily driven by unobserved selection into subsidy use.

[30] We will conduct robustness tests for the full set of outcomes when data access is renewed.

6 Firm Learning

6.1 Empirical approach

The previous Section documented that Bonus Donne firms increased recruitment from the targeted group, without any decline in firm labor productivity. This raises a key question: if these workers are of comparable quality, why were they not hired before the subsidy? A possible explanation, introduced in Section 4, is that firms believe women with long employment interruptions to be poor matches, leaving them in a learning trap. By lowering the effective cost of recruiting from this group, the subsidy encouraged firms to experiment with these workers and learn about their match quality. As our model predicts, the adjustment in firms' beliefs depends on their initial experience with Bonus Donne hires: firms that receive more positive signals about match quality with the target group may update their beliefs upwards and, consequently, hire from this group in the future.

To explore this learning mechanism empirically, we examine whether variation in firms' initial experiences with Bonus Donne hires affects their subsequent hiring behavior. We focus on treated firms that first use the subsidy early (treatment cohorts 2013 to 2016—a total of 6,428 firms) in order to track their hiring patterns over time following their initial use of the subsidy. We ask whether firms that hired better matches in the year of their first subsidized hire were subsequently more likely to hire again from the target group, relative to treated firms whose initial hires were perceived to be poorer matches. To do so, we classify these early-treated firms based on the quality of their initial match with Bonus Donne hires and distinguish between those hiring a "good" versus a "bad" match—that is, a worker perceived to be more more versus less productive for the firm.

To construct a proxy for firm-match quality for each Bonus Donne hire, we use individual-level data to estimate a Mincerian wage regression that relates (log) weekly wages to key observable worker characteristics (age, contract type, occupation, hiring cohort) and firm characteristics (location, 2-digit industry, total employment). The estimated residual—the portion of the hiring wage unexplained by these characteristics—is intended to proxy firm-specific match quality

at the time of hiring, in line with our conceptual framework in Section 4.[31] We estimate the Mincerian regression on the sample of initial hires—women hired for the first time under the Bonus Donne between 2013 and 2016—restricting the sample to the years in which these hires occur.[32] We then discretize the distribution of residuals and classify Bonus Donne hires below the median as bad matches and those above the median as good matches.[33] Treated firms are then classified based on the quality of their initial (time 0) Bonus Donne hire(s): firms hiring an above-median worker are labeled as "good-match" firms, while those hiring a below-median worker are labeled "bad-match" firms.[34]

The empirical analysis employs a difference-in-differences framework comparing hiring patterns at bad-match and good-match firms in the years following the first subsidized hire. The year of first subsidy use—time 0, which is used to classify firms based on match quality—is therefore excluded from the estimation sample. We estimate the following specification:

$$y_{jt} = \alpha + \gamma_1 \cdot Post + \gamma_2 \cdot Post * Good_j + \xi_j + \epsilon_{jt} \tag{3}$$

where $Good_j$ is an indicator variable equal to one for good-match treated firms and zero for bad-match treated firms. The variable $Post$ takes value one for years 1 to 5 after the first subsidized hire, and zero for the years -4 to -1. Firm fixed effects (ξ_j) control for time-invariant differences across firms, such as workforce composition or firm policies. The coefficient of interest, γ_2, captures the differential change in outcomes between good- and bad-match firms, averaged over the post period, relative to the pre-treatment baseline. By construction, this comparison excludes the event year (time 0)—which coincides with the first subsidized hire and therefore features a sharp increase in subsidized hiring by definition (see Section 5)—and

[31] Studies in the U.S. retail sector, where turnover is high, often use worker tenure or turnover decisions as productivity measures (Autor & Scarborough, 2008; Benson & Lepage, 2025). Such ex-post measures are less suitable in the Italian context, where labor markets are less flexible and turnover is less frequent. Using tenure as a proxy for worker productivity would mechanically lower subsequent subsidy use and hiring for firms with higher retention, introducing negative bias. Our approach instead leverages detailed administrative records on wages and worker characteristics to construct a cleaner proxy for firms' perceived match quality *at the time of hiring*, as reflected in entry wages. This measure corresponds closely to the wage-setting assumption in our theoretical framework (Section 4).

[32] We run a Mincer wage regression pooling together all subsidized female hires at Bonus Donne firms between 2013 and 2016. In addition to the controls listed in the main text, the regression also includes year fixed effects.

[33] The empirical results outlined are robust to alternative classification using the first and third tercile.

[34] If a firm hires multiple Bonus Donne workers at time 0, it is classified based on the modal worker.

instead isolates how initial experiences with subsidized hires shape subsequent hiring behavior. Standard errors are clustered at the firm level.

In the baseline specification, identification comes from differences between treated (Bonus Donne) firms with good versus bad initial matches. This design relies on parallel trends between these two groups. To further corroborate this approach and allow for differential pre-trends between good- and bad-match firms, we extend the analysis by incorporating the matched control group described in Section 5. Specifically, we estimate a triple-differences specification that compares: (i) good-match treated firms to their matched controls, (ii) bad-match treated firms to their matched controls, and (iii) the difference between these two effects, thereby differencing out any common shocks affecting treated firms.[35]

6.2 Results

Table 3 presents estimates of the γ_2 coefficient in Equation 3 for various firm outcomes. The results show that hiring patterns following a firm's first subsidized hire under the Bonus Donne depend on its initial experience with the program. Good-match firms, which hired a better match through the subsidy, hire 2% more Bonus Donne workers in later periods compared to bad-match firms (Column 1). Good-match firms also increase their hiring of workers from the targeted group, as hiring of long-term non-employed women and mothers rises by 2.3% and 1.7%, respectively (Columns 2-3)—magnitudes comparable to the overall post-event increase in hiring from these groups at baseline (Figure 2 Panels (b)-(c)), indicating that the within-treated differences are economically meaningful relative to the baseline response. This pattern is consistent with firms updating their beliefs about match quality for the targeted group, in line with the learning mechanism outlined in our conceptual framework. Finally, Column 4 shows no significant effect on total firm size. This suggests that the increased hiring from the targeted group reflects a change in hiring composition rather than differential firm growth between good-

[35]The estimated model is as follows, where the coefficient of interest γ_4 captures the triple-differences effect and indicator variables $Post$, $Good_j$, and T_j are defined in the main text:

$$y_{jt} = \alpha + \gamma_1 \cdot Post + \gamma_2 \cdot Post * Good_j + \gamma_3 \cdot Post * T_j + \gamma_4 \cdot Post * T_j * Good_j + \xi_j + \epsilon_{jt} \quad (4)$$

and bad-match firms.

We provide two robustness checks. First, Appendix Table A.4 confirms our findings when good- and bad-match firms are defined based on whether their (modal) Bonus Donne hire lies in the first or third tercile of the distribution of Mincer residuals, rather than above versus below the median. Second, Appendix Table A.5 confirms the baseline results under the triple-differences design, where good- and bad-match treated firms are also compared with their matched controls. Notably, we find no statistically significant differences in any of the outcomes between the control matched to good- and bad-match treated firms (not reported). This is reassuring, as it suggests that the different hiring trajectories between good- and bad-match firms are unlikely to be driven by differential shocks, but instead reflect differences in their initial exposure to the targeted group. Taken together, these results indicate that early positive experience with workers from the targeted group increases firms' willingness to hire from this group in subsequent periods, consistent with our theoretical framework and models of experiential learning (Benson & Lepage, 2025; Lepage, 2024; Leung, 2018).

7 Worker Outcomes

7.1 Empirical approach

So far, we have focused on how firms responded to the Bonus Donne. This final section turns to the workers themselves and examines their long-term labor market outcomes.

Worker matching. To examine worker-level effects, we build on the firm-level design by matching each Bonus Donne hire to a comparable control worker. We focus on female workers hired from non-employment at treated and control firms, using the matched firm sample described in Section 5 and restricting attention to the first year treated firms hire a Bonus Donne worker (between 2013 and 2019). First, we identify subsidized hires at treated firms in the initial year of subsidy uptake. Each of these workers is matched to a newly hired female worker in the same year at matched control firms, with similar observable characteristics. The worker matching is based on a propensity-score logit model for the probability that a worker is hired through

the subsidy, controlling for worker age, contract type (fixed-term vs. open-ended, part-time vs. full-time), broad occupational categories (manager, white collar, blue collar), and the length of the preceding non-employment spell. This approach compares the careers of observationally similar female workers—of similar age, hired after comparable non-employment spells, into the same occupation and contract type, and by firms on similar trajectories—who differ only in whether they were hired with or without the subsidy. Crucially, our administrative data allow us to follow workers not only at the hiring firm but also at subsequent employers. Compared to the firm-level analysis, this design relies on weaker assumptions for internal validity, as a firm's decision to use the subsidy when hiring a worker is unlikely to be driven by individual workers' choices. This, in turn, allows us to more cleanly isolate the effect of being hired under the subsidy on workers' careers. Unlike the firm-level analysis, which reflects changes in hiring behavior and firm growth following subsidy adoption, the worker-level design compares similar hires and isolates the effect of subsidy exposure on workers' careers.

Our matched sample consists of 13,336 Bonus Donne workers, and an equal number of matched control workers. Appendix Table A.6 confirms that the two groups share similar characteristics.

Worker event-study design. Similarly to the firm analysis, we estimate a staggered event-study design on the matched sample of new hires:

$$y_{it} = \phi_i + \rho_t + \sum_{k=-4}^{5} \gamma_k \cdot \mathbb{1}[t = t_i^* + k] + \sum_{k=-4}^{5} \beta_k \cdot \mathbb{1}[t = t_i^* + k] \cdot T_i + e_{it}, \tag{5}$$

where y_{it} denotes the outcome variable for worker i in calendar year t, T_i is a treatment indicator equal to one for Bonus Donne workers and zero for the matched control group, and $t_i^* \in [2013, \ldots, 2019]$ is the year treated worker i was hired through the Bonus Donne (for a control worker, this is the hiring year of their matched treated worker). The specification controls for time-invariant, worker-specific unobserved characteristics through worker fixed effects ϕ_i, for aggregate shocks through year fixed effects ρ_t, and for time-to-treatment effects through the γ_k coefficients. The coefficients of interest are the β_k's, denoting the difference in outcomes between treated and control workers relative to the year before being hired under the Bonus

Donne (β_{-1} normalized to zero). They identify the causal effect of being hired under the subsidy on workers' careers, under a parallel trends assumption: in the counterfactual scenario where Bonus Donne hires would have been hired without the subsidy, their careers would have evolved similarly to those of their matched controls. Standard errors are clustered at the worker level.

Our outcomes of interest are wages and employment trajectories both at the original hiring firm and in the broader labor market. By comparing a Bonus Donne hire with an observationally similar worker who was not hired through the subsidy, the design allows us to address two central questions. First, by accounting for worker selection, we directly assess whether the hiring subsidy is passed through to workers in the form of higher wages, or retained by firms. Second, by tracking workers for up to five years after hiring, we examine whether the Bonus Donne improved long-term labor market attachment even after the subsidy expires—a central goal of hiring subsidies. Because the subsidy is extended in duration when a worker is hired on, or converted to, a permanent contract, it creates an incentive for firms to retain subsidized hires and may also enhance workers' attractiveness to other employers through accumulated experience.

7.2 Results

Figure 4, Panel (a), shows event-study coefficients for weekly wages net of employer social security contributions, that is, workers' take-home pay. This outcome is measured conditional on employment (i.e., wages are set to missing if the worker is not in employment) and was not targeted in our matching algorithm. Before being hired, the trends of net wages of treated and control workers were not significantly different. At time 0, the year treated workers were hired through the Bonus Donne and matched control workers were also hired, we observe no divergence in take-home wages, and the same holds in subsequent periods. This result provides evidence that there is no pass-through of the hiring subsidy to workers in the form of higher wages, in line with Rubolino (2025) and Jain, Mommaerts, and Weaver (2025).

We next examine the effects on workers' employment trajectories. Figure 4 Panel (b) shows that one year post-hire, subsidized workers are about 10 percentage points more likely to be retained by their initial employer (dashed lines). This likely reflects the financial incentives offered to

firms by the Bonus Donne rules (Section 2.2). While the retention gap narrows over time, it does not sharply decline by year 2—when the subsidy expires—nor does it disappear in the long run: even five years later, subsidized hires remain nearly 5 percentage points more likely to be employed at their original firm compared to control workers. The findings suggest that firms do not churn workers after the subsidy expires, but tend to retain them in the long run.

We also investigate effects on workers' employment at any firm. Overall labor market attachment may differ between Bonus Donne hires and matched control workers, not only because of direct retention incentives for firms, but also through indirect effects. For example, longer initial employment spells of Bonus Donne hires may improve their skills or signal higher productivity to other employers (Barham, Cadena, & Turner, 2023). As shown in Figure 4 Panel (b) (solid lines), treated workers are 8 percentage points more likely to be employed than control peers one year after hire, with the effect persisting in subsequent years. While nearly half of this difference is accounted for by higher retention at the original firm, the remainder reflects a greater likelihood of treated workers transitioning to other firms. These findings suggest that hiring subsidies can enhance labor market attachment for the targeted group—not only through immediate job retention, but also through broader employment gains, in contrast to some earlier evidence from temporary work programs (Barham, Cadena, & Turner, 2023).

8 Conclusion

Increasing women's labor market participation is a key policy priority in many countries. While most policy interventions have focused on supply-side measures such as parental leave or childcare provision, less attention has been paid to the role of employer demand in shaping women's employment opportunities. This paper studies whether a hiring subsidy introduced in Italy and targeted at women with extended employment interruptions can improve women's labor market outcomes by encouraging firms to hire from this group.

Leveraging administrative employer-employee data that record firms' take-up of the subsidy and allow us to track both firms and workers over time, we document several key findings. First, firms using the subsidy diversify their hiring by recruiting more women with long career

interruptions, including mothers. While the effect is strongest in the year of take-up, hiring patterns do not revert to pre-treatment levels thereafter. Treated firms continue to recruit more women with extended employment interruptions and more mothers relative to matched control firms. Importantly, these changes do not come at the expense of firm performance, as treated firms experience no decline in productivity. This suggests that encouraging firms to broaden their hiring toward underrepresented groups need not involve a trade-off with firm performance.

Second, we present evidence suggesting that firms learn from their initial matches with subsidized workers. Firms that experience better initial matches with subsidized workers are significantly more likely to continue hiring from the targeted group in subsequent years. This pattern is consistent with the employer learning mechanism formalized in our conceptual framework: by lowering the cost of experimentation, hiring subsidies generate information about match quality between firms and workers in the targeted group, and can shift employers' beliefs about the productivity of workers with extended employment interruptions.

Third, we show that women hired under the subsidy experience persistent employment gains. Even five years after hiring, subsidized workers are substantially more likely to be employed relative to comparable women hired at control firms. Importantly, this effect reflects not only higher retention at the original firm but also stronger attachment to the broader labor market. Given evidence that career breaks following childbirth can reduce re-entry probabilities (Adda, Dustmann, & Stevens, 2017; Kleven, Landais, & Sogaard, 2019), these findings suggest that hiring subsidies may help reconnect mothers to stable employment trajectories. This contrasts with Barham, Cadena, and Turner (2023), who find that long-run employment gains from a subsidized work program in Colorado operate primarily through retention at the initial employer. Differences in the target population and the institutional context may help explain why employment gains in our setting extend beyond the initial hiring firm. Finally, we find no evidence that the subsidy was passed through to workers in the form of higher wages.

These findings highlight the role that demand-side interventions directly affecting firms' hiring decisions can play in addressing gender inequalities in the labor market. Specifically, hiring subsidies may help address barriers that arise when firms are uncertain about the productivity of

workers with extended employment interruptions, and may therefore be effective when targeted at groups facing such hiring barriers. The lack of productivity losses also alleviates concerns that such hiring might come at the expense of firm performance. Furthermore, the Bonus Donne experience suggests that linking subsidy generosity to stable employment relationships may foster stable attachment to the labor market for targeted workers.

Finally, our analysis suggests that temporary exposure to the policy may lead to lasting shifts in firms' hiring practices, beyond the initial subsidy adoption. This relates to a broader literature showing that transitory policies can generate persistent changes in firm behavior, for instance through firms' investment in screening capital (Miller, 2017). In our framework, this occurs because hiring subsidies encourage firms to experiment and learn about the productivity of workers with extended employment interruptions, thus expanding employment opportunities for women who face difficulties re-entering the labor market.

References

Adda, J., Dustmann, C., & Stevens, K. (2017). The career costs of children. *Journal of Political Economy*, *125*(2), 293–337.

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, *30*(2), 175–187.

Albanesi, S., Olivetti, C., & Petrongolo, B. (2023). Families, labor markets, and policy. In *Handbook of the economics of the family* (pp. 255–326, Vol. 1). Elsevier.

Amano-Patiño, N., Contractor, Z., & Aramburu, J. (2025, March). *Is affirmative action in employment still effective in the 21st century?* [Unpublished manuscript].

Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, *34*(3), 545–579.

Arrow, K. J. (1973). The theory of discrimination (O. Ashenfelter & A. Rees, Eds.), 3–33.

Arulampalam, W. (2001). Is unemployment really scarring? effects of unemployment experiences on wages. *The Economic Journal*, *111*(475), 585–606.

Autor, D. H., & Scarborough, D. (2008). Does job testing harm minority workers? Evidence from retail establishments. *The Quarterly Journal of Economics*, *123*(1), 219–277.

Baker, A. C., Larcker, D. F., & Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, *144*(2), 370–395.

Barham, T., Cadena, B. C., & Turner, P. S. (2023). Taking a chance on workers: Evidence on the effects and mechanisms of subsidized employment from an rct. *IZA Discussion Papers*, *16221*.

Benson, A., Board, S., & Vehn, M. M.-t. (2024). Discrimination in hiring: Evidence from retail. *The Review of Economic Studies*, *91*(4), 1956–1987.

Benson, A., & Lepage, L.-P. (2025). *Learning to discriminate on the job* (RFBerlin Discussion Paper Series No. 25116). ROCKWOOL Foundation Berlin (RFBerlin).

Benzarti, Y., & Harju, J. (2021a). Can payroll tax cuts help firms during recessions? *Journal of Public Economics*, *200*, 104472.

Benzarti, Y., & Harju, J. (2021b). Using payroll tax variation to unpack the black box of firm-level production. *Journal of the European Economic Association*, *19*(5), 2737–2764.

Bertrand, M., Black, S. E., Jensen, S., & Lleras-Muney, A. (2018). Breaking the glass ceiling? the effect of board quotas on female labour market outcomes in norway. *The Review of Economic Studies*, *86*(1), 191–239.

Bíró, A., Branyiczki, R., Lindner, A., Márk, L., & Prinz, D. (2024). *Firm heterogeneity and the impact of payroll taxes* [Unpublished manuscript, March 2024 draft].

Bishop, J. (1981). Employment in construction and distribution industries: The impact of the new jobs tax credit. In *Studies in labor markets* (pp. 209–246). University of Chicago Press.

Boeri, T. (2011). Chapter 13 - institutional reforms and dualism in european labor markets. In D. Card & O. Ashenfelter (Eds.). Elsevier.

Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review*, *109*(10), 3395–3436.

Cahuc, P., Carcillo, S., & Le Barbanchon, T. (2019). The effectiveness of hiring credits. *The Review of Economic Studies*, *86*(2), 593–626.

Campa, P., Casarico, A., & Profeta, P. (2010). Gender culture and gender gap in employment. *CESifo Economic Studies*, *57*(1), 156–182.

Carta, F., De Philippis, M., Rizzica, L., & Viviano, E. (2023). *Women, labour markets and economic growth* (Workshops and Conference No. 26). Bank of Italy.

Casarico, A., & Lattanzio, S. (2023). Behind the child penalty: Understanding what contributes to the labour market costs of motherhood. *Journal of Population Economics*, *36*(3), 1489–1511.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, *134*(3), 1405–1454.

Coate, S., & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *American Economic Review*, *83*(5), 1220–1240.

Cohen, J., Johnston, A. C., & Lindner, A. (2025). Skill depreciation during unemployment: Evidence from panel data. *American Economic Journal: Applied Economics*, *17*(3), 208–235.

Cornell, B., & Welch, I. (1996). Culture, information, and screening discrimination. *Journal of Political Economy*, *104*(3), 542–571.

Correll, S. J., Benard, S., & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, *112*(5), 1297–1338.

Cullen, Z., Dobbie, W., & Hoffman, M. (2022). Increasing the demand for workers with a criminal record. *The Quarterly Journal of Economics*, *138*(1), 103–150.

Eriksson, S., & Rooth, D.-O. (2014). Do employers use unemployment as a sorting criterion when hiring? evidence from a field experiment. *American Economic Review*, *104*(3), 1014–1039.

Eurostat. (2025). *Employment - annual statistics: Employment of men and women by age groups* [Statistics Explained, Accessed: October 2025]. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Employment_-_annual_statistics#Employment_of_men_and_women_by_age_groups

Fryer, R. G. J., & Loury, G. C. (2013). Valuing diversity. *Journal of Political Economy*, *121*(4), 747–774.

Giavazzi, F., Serafinelli, M., & Schiantarelli, F. (2013). Attitudes, policies, and work. *Journal of the European Economic Association*, *11*(6), 1256–1289.

Hamersma, S. (2003). The work opportunity and welfare-to-work tax credits: Participation rates among eligible workers. *National Tax Journal*, *56*(4), 725–38.

Huttunen, K., Pirttilä, J., & Uusitalo, R. (2013). The employment effects of low-wage subsidies. *Journal of Public Economics*, *97*(1), 49–60.

Jain, M., Mommaerts, C., & Weaver, J. (2025). *Hiring subsidies for the disadvantaged: Evidence from the work opportunity tax credit* [Unpublished manuscript].

Katz, L. (1998). Wage subsidies for the disadvantaged. Russell Sage Foundation.

Kleven, H., Landais, C., & Leite-Mariante, G. (2024). The child penalty atlas. *The Review of Economic Studies*, *92*(5), 3174–3207.

Kleven, H., Landais, C., & Sogaard, J. E. (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics*, *11*(4), 181–209.

Kroft, K., Lange, F., & Notowidigdo, M. J. (2013). Duration dependence and labor market conditions: Evidence from a field experiment. *The Quarterly Journal of Economics*, *128*(3), 1123–1167.

Ku, H., Schönberg, U., & Schreiner, R. C. (2020). Do place-based tax incentives create jobs? *Journal of Public Economics*, *191*, 104105.

Kuhn, P., & Yu, L. (2021). How costly is turnover? evidence from retail. *Journal of Labor Economics*, *39*(2), 461–496.

La Repubblica. (2011). *Monti al senato per la fiducia il testo integrale del discorso*. https://www.repubblica.it/politica/2011/11/17/news/monti_al_senato_per_la_fiducia_il_testo_integrale_del_discorso-25168289/

Lalive, R., Schlosser, A., Steinhauer, A., & Zweimüller, J. (2014). Parental leave and mothers' careers: The relative importance of job protection and cash benefits. *Review of Economic Studies*, *81*(1), 219–265.

Lepage, L. P. (2024). Experience-based discrimination. *American Economic Journal: Applied Economics*, *16*(4), 288–321.

Leung, M. D. (2018). Learning to hire? hiring as a dynamic experiential learning process in an online market for contract labor. *Management Science*, *64*(12), 5651–5668.

Li, D., Raymond, L., & Bergman, P. (2025). Hiring as exploration. *The Review of Economic Studies*, *93*(2), 1200–1240.

Lundberg, S. J., & Startz, R. (1983). Private discrimination and social intervention in competitive labor markets. *American Economic Review*, *73*(3), 340–347.

Maida, A., & Weber, A. (2022). Female leadership and gender gap within firms: Evidence from an italian board reform. *ILR Review*, *75*(2), 488–515.

Miller, A., & Segal, C. (2012). Does temporary affirmative action produce persistent effects? evidence from the federal contractor program. *Review of Economics and Statistics*, *94*(3), 611–627.

Miller, C. (2017). The persistent effect of temporary affirmative action. *American Economic Journal: Applied Economics*, *9*(3), 152–90.

Miller, C., Peck, J., & Seflek, M. (2022). Nitaqat, missing women, integration costs, and big push policies in the saudi labor market. *American Economic Journal: Applied Economics*, *14*(2), 51–77.

Neumark, D. (2013). Spurring job creation in response to severe recessions: Reconsidering hiring credits. *Journal of Policy Analysis and Management*, *32*, 142–71.

Neumark, D., & Grijalva, D. (2017). The employment effects of state hiring credits. *ILR Review*, *70*(5), 1111–1145.

Olivetti, C., & Petrongolo, B. (2008). Unequal pay or unequal employment? a cross-country analysis of gender gaps. *Journal of Labor Economics*, *26*(4), 621–654.

Perloff, J. M., & Wachter, M. L. (1979). The new jobs tax credit: An evaluation of the 1977-78 wage subsidy program. *The American Economic Review*, *69*(2), 173–179.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review*, *62*(4), 659–661.

Rubolino, E. (2025). *Taxing or subsidizing the gender gap? labor market responses to a female hiring subsidy* [Unpublished manuscript; earlier version posted July 19, 2021, last revised October 29, 2025].

Saez, E., Matsaganis, M., & Tsakloglou, P. (2012). Earnings determination and taxes: Evidence from a cohort-based payroll tax reform in greece. *The Quarterly Journal of Economics*, *127*(1), 493–533.

Saez, E., Schoefer, B., & Seim, D. (2019). Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in sweden. *American Economic Review*, *109*(5), 1717–1763.

Saez, E., Schoefer, B., & Seim, D. (2021). Hysteresis from employer subsidies. *Journal of Public Economics*, *200*, 104459.

Schönberg, U., & Ludsteck, J. (2014). Expansions in maternity leave coverage and mothersâ labor market outcomes after childbirth. *Journal of Labor Economics*, *32*(3), 469–505.

Sun, L. (2021). Eventstudyinteract: Stata module to implement the interaction weighted estimator for an event study.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199.

Weisshaar, K. (2018). From opt out to blocked out: The challenges for labor market re-entry after family-related employment lapses. *American Sociological Review*, *83*(1), 34–60.

Tables and Figures

Tables

Table 1: Firm Descriptive Statistics

	(1) Bonus Donne Firm	(2) Average Italian Firm
Number of employees	48.17	17.96
	(910.89)	(263.47)
Female share (employment)	0.57	0.44
	(0.33)	(0.36)
Share full-time contract (employment)	0.63	0.70
	(0.34)	(0.34)
Share open-ended contract (employment)	0.80	0.83
	(0.27)	(0.28)
Female share (new hires)	0.54	0.42
	(0.40)	(0.41)
Number of new hires 2010-12	12.32	3.47
	(350.99)	(75.42)
Mean ln weekly wage	5.84	5.87
	(0.34)	(0.41)
Mean ln weekly wage (women)	5.82	5.85
	(0.36)	(0.46)
Value added per worker	35.10	40.42
	(32.97)	(58.21)
Number of firms	26,497	660,764

Notes: The table compares mean values for firms that have used the Bonus Donne for at least one female hire (Column 1) to other Italian firms that have never made use of the subsidy (Column 2). Data as of 2012. The number of new hires 2010-12 is the mean annual number of new hires at the firm, averaged between 2010 and 2012. Ln weekly wage is the natural logarithm of the weekly wage earned by workers at the firm, averaged across workers. Standard deviations in parentheses.

Table 2: Balancing Table, Firms

	(1) Treated firms	(2) Control firms
Targeted for matching		
Mean ln weekly wage	5.94	5.94
	(0.28)	(0.37)
Mean ln weekly wage (women)	5.92	5.91
	(0.30)	(0.40)
Number of employees	44.89	43.83
	(512.53)	(594.65)
Female share (employment)	0.55	0.55
	(0.32)	(0.32)
Not targeted for matching		
Share full-time contract (employment)	0.60	0.63
	(0.35)	(0.34)
Share open-ended contract (employment)	0.83	0.77
	(0.24)	(0.30)
Number of firms	19,135	19,135

Notes: The table shows mean values for treated (Column 1) and control (Column 2) firms in our matched sample. The matching is implemented using propensity score matching based on firm-level characteristics measured in the three years preceding the treated firm's first use of the subsidy: average log weekly wage, log total employment, female employment share, and quartiles of the number of hires. In addition, control firms are required to hire at least one female worker in the same calendar year in which their matched treated firm first hires under the Bonus Donne (see Section 5 for details). Treated firms are those that have used the Bonus Donne at least once. Ln weekly wage is the natural logarithm of the weekly wage earned by workers at the firm, averaged across workers. Descriptive statistics are computed in the period before the adoption of the subsidy. Standard deviations in parentheses.

Table 3: Good-Match versus Bad-Match Firms

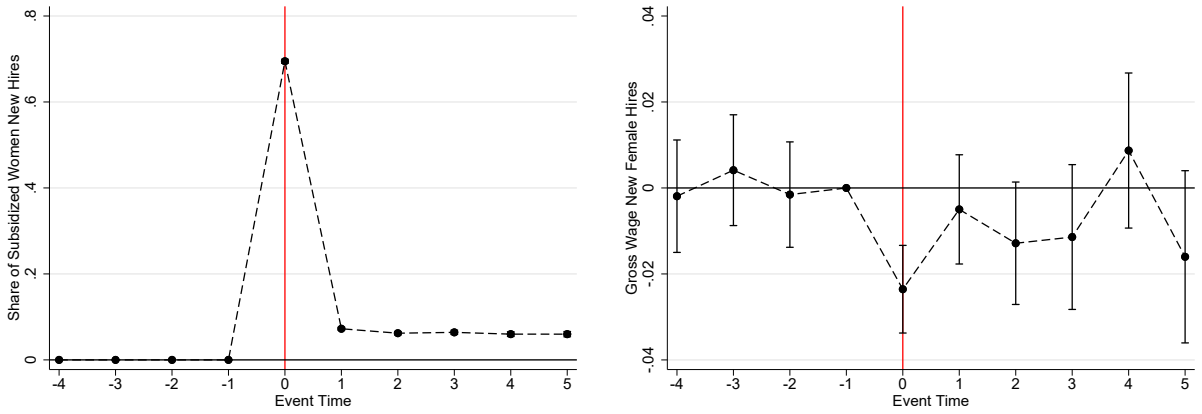
	(1) Ln Subsidized Female Hires	(2) Ln New Hires From Long-Term Non-Employment	(3) Ln Mothers New Hires	(4) Ln Total Firm Size
Estimate	0.0211**	0.0226***	0.0171***	0.0055
	(0.0094)	(0.0075)	(0.0064)	(0.0121)
Observations	70,224	70,224	70,224	70,224
R-squared	0.342	0.721	0.646	0.924

Notes: The table reports the estimated γ_2 coefficient from Equation 3, denoting the differential effect between treated firms hiring a more versus less productive worker through the Bonus Donne (good-match versus bad-match firms). Good (bad) match firms are those where the modal Bonus Donne hire has above (below) median match productivity. Match productivity computed as the residual of a Mincer wage regression (see Section 6 for details). The coefficient is estimated in the post-treatment periods 1 to 5, thus excluding the treatment take-up period. Column (1): The outcome is the natural logarithm of (1+) the number of female workers hired under the Bonus Donne; Column (2): The outcome is the natural logarithm of (1+) the number of female new hires from long-term ($\geq$ 3 years) non-employment; Column (3): The outcome is the natural logarithm of (1+) the number of newly hired mothers at the firm (mothers defined as workers that had maternity leave in the previous four years); Column (4): The outcome is the natural logarithm of the total number of employees at the firm. Standard errors clustered at the firm level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figures

Figure 1: Share and Cost of Female Workers Hired Through the Bonus Donne

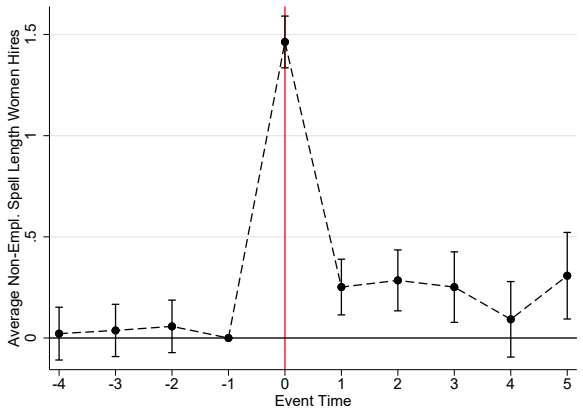
(a) Share Hired Under Bonus Donne (b) Mean Log Weekly Wages, Gross of Payroll Taxes



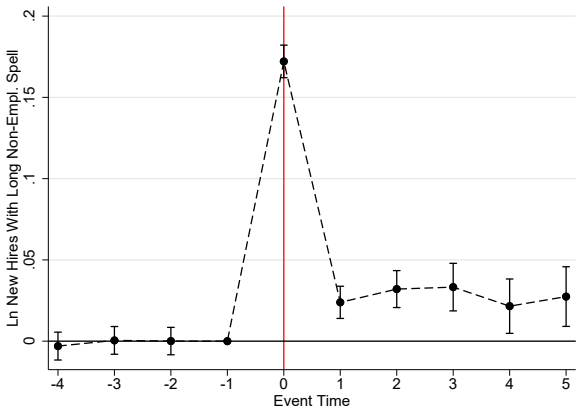
Notes: The figure plots coefficient estimates from Equation 2 denoting differential effects for treated firms using the Bonus Donne versus matched control firms, before and after subsidy adoption (event time = 0). Panel (a): The outcome is female workers hired through the subsidy as a share of new female hires; Panel (b): The outcome is the total labor cost for new female hires at the firm each period. The red vertical line denotes the period when treated firms hire at least one worker through the Bonus Donne for the first time. The solid vertical lines denote 95% confidence intervals. The underlying event study regression control for firm, year and time from-event fixed-effects. Standard errors clustered at the firm level. Matching between treated and control firms is implemented using propensity score matching based on firm-level characteristics measured in the three years preceding the treated firm's first use of the subsidy: average log weekly wage, log total employment, female employment share, and quartiles of the number of hires. In addition, control firms are required to hire at least one female worker in the same calendar year in which their matched treated firm first hires under the Bonus Donne (see Section 5 for details).

Figure 2: Effects on Firm Hiring Patterns Towards Target Group

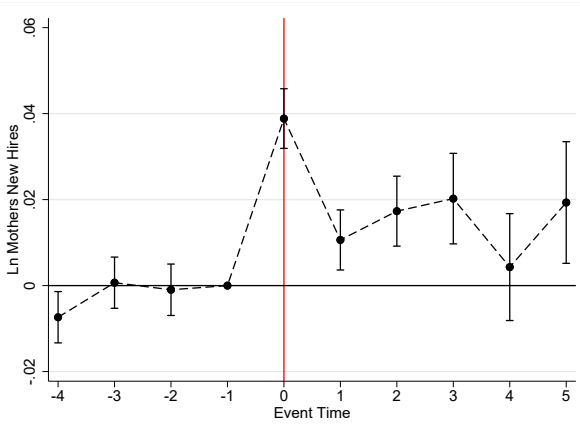
(a) Previous Non-Employment Period



(b) Female Hires with Long Non-Employment Spells

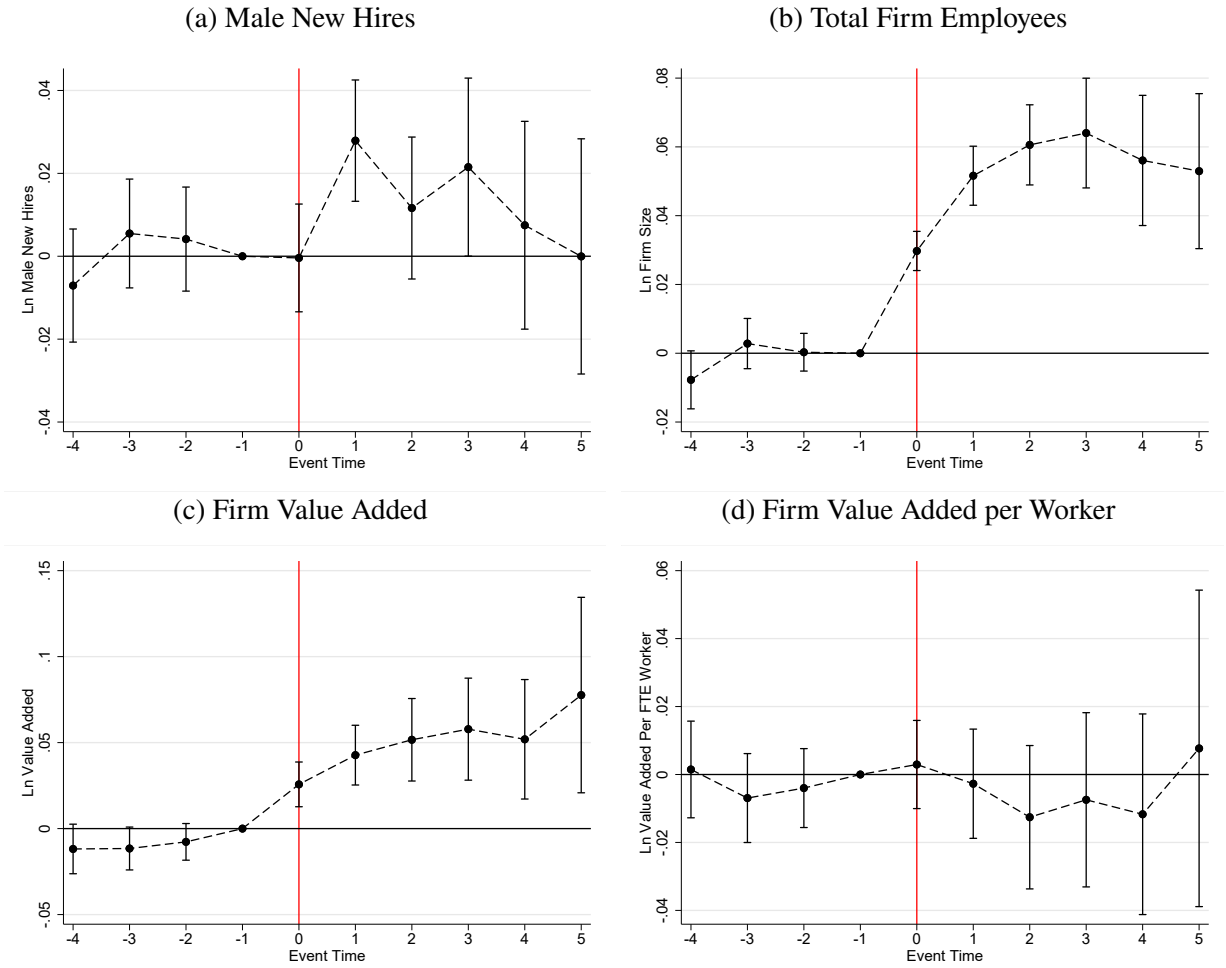


(c) Mothers



Notes: The figure plots coefficient estimates from Equation 2 denoting differential effects for treated firms using the Bonus Donne versus matched control firms, before and after subsidy adoption (event time = 0). Panel (a): The outcome is the average length of non-employment (in years) for new female hires; Panel (b): The outcome is the natural logarithm of (1+) the number of female new hires from long-term ($\geq$ 3 years) non-employment; Panel (c): The outcome is the natural logarithm of (1+) the number of newly hired mothers at the firm (mothers defined as workers that had maternity leave in the previous four years). The red vertical line denotes the period when treated firms hire at least one worker through the Bonus Donne for the first time. The solid vertical lines denote 95% confidence intervals. The underlying event study regression control for firm, year and time from-event fixed-effects. Standard errors clustered at the firm level. Matching between treated and control firms is implemented using propensity score matching based on firm-level characteristics measured in the three years preceding the treated firm's first use of the subsidy: average log weekly wage, log total employment, female employment share, and quartiles of the number of hires. In addition, control firms are required to hire at least one female worker in the same calendar year in which their matched treated firm first hires under the Bonus Donne (see Section 5 for details).

Figure 3: Other Firm Outcomes

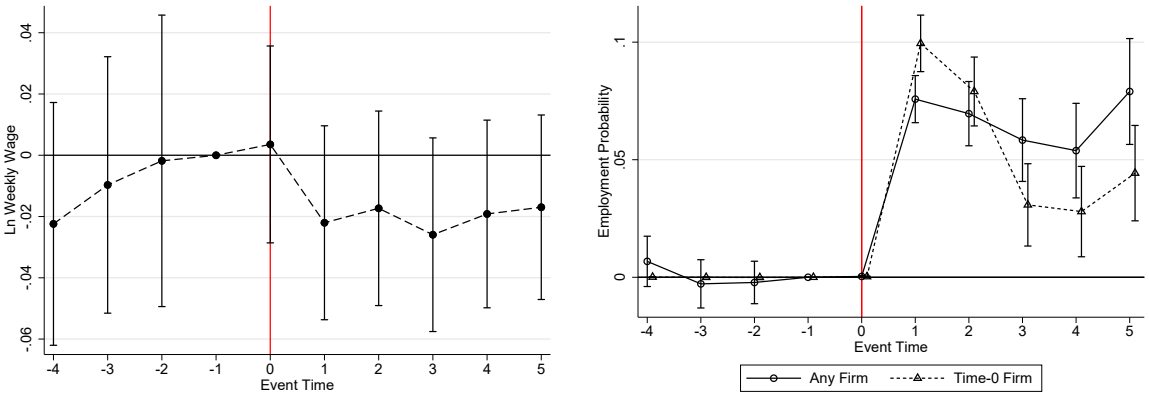


Notes: The figure plots coefficient estimates from Equation 2 denoting differential effects for treated firms using the Bonus Donne versus matched control firms, before and after subsidy adoption (event time = 0). Panel (a): The outcome is the natural logarithm of (1+) the number of men among new hires; Panel (b): The outcome is the natural logarithm of the total number of employees at the firm; Panel (c): The outcome is the natural logarithm of firm value added; Panel (d): The outcome is the natural logarithm of firm value added per full-time equivalent worker. The red vertical line denotes the period when treated firms hire at least one worker through the Bonus Donne for the first time. The solid vertical lines denote 95% confidence intervals. The underlying event study regression control for firm, year and time from-event fixed-effects. Standard errors clustered at the firm level. Matching between treated and control firms is implemented using propensity score matching based on firm-level characteristics measured in the three years preceding the treated firm's first use of the subsidy: average log weekly wage, log total employment, female employment share, and quartiles of the number of hires. In addition, control firms are required to hire at least one female worker in the same calendar year in which their matched treated firm first hires under the Bonus Donne (see Section 5 for details).

Figure 4: Worker-Level Effects: Net Wages and Employment

(a) Net Wages

(b) Employment Probabilities



Notes: The figure plots coefficient estimates from Equation 5 denoting differential effects for treated workers hired through the Bonus Donne versus matched control workers, before and after the hiring year (event time = 0). Treated and control workers are obtained from the matched sample of firms used for the firm-level analysis, see Section 7 for details. Panel (a): The outcome is the natural logarithm of the worker's net weekly wage (missing if the worker is non-employed); Panel (b): The outcome is worker's employment probability at any firm (solid lines) and at the time-0 firm (dashed lines). The red vertical line denotes the period when treated workers are hired through the Bonus Donne. Vertical solid lines denote 95% confidence intervals. The underlying event study regression control for worker, year and time-to-treatment fixed-effects. Standard errors clustered at the firm level. Matching between treated and control workers is based on the sample female workers hired from non-employment at treated and control firms, using the matched firm sample described in Section 5 and restricting attention to the first year treated firms adopt the subsidy (between 2013 and 2019). We identify female workers hired through the Bonus Donne at treated firms in the initial year of subsidy uptake. Each of these workers is matched to a newly hired female worker in the same year at matched control firms. Matching is based on a propensity-score logit model for the probability that a worker is hired through the subsidy, controlling for worker age, contract type (fixed-term vs. open-ended, part-time vs. full-time), broad occupational categories (manager, white collar, blue collar), and the length of the preceding non-employment spell (see Section 7 for details).

Appendices

A.1 Policy Background and Political Setting

Following the Great Financial Crisis, many European countries were struck by the Sovereign Debt Crisis in the early 2010s. These included Italy, where turmoil in financial markets led to the fall of the incumbent government in November 2011. A technocratic government led by Prime Minister Mario Monti came to power with the aim of implementing structural reforms to bring public debt back under control and promote economic growth. These reforms were implemented in short time, due to pressures from international investors and supranational institutions. Two key labor market reforms were introduced six months apart: a pension reform, approved on December 22, 2011, and a labor market reform, approved on June 28, 2012 (Law 92/2012, which also introduced the Bonus Donne). The former created new pension rules to ease pressure on public debt.[36] The latter had four major components: *i*) it eased employment protection legislation for workers under an open-ended contract by relaxing rules on economic dismissal (a very controversial policy decision); *ii*) it made temporary and atypical contracts marginally more stringent;[37] *iii*) it expanded the coverage of the unemployment insurance scheme; and *iv*) it introduced the hiring subsidy which is the focus of our paper—the Bonus Donne. Importantly for our empirical design, the latter was the only subsidy targeted to firms and to the hiring of women.

There was strikingly little news coverage of the Bonus Donne. In contrast, most media attention focused on the pension reform and on the easing of the employment protection legislation. These policy items were highly controversial in the public debate, as either worsening pension conditions for all cohorts or allegedly going against some of the principles laid out by the Italian constitution (Article 18) by allowing the firing of workers for economic reasons. Google Trends searches clearly illustrate that these policies attracted far more media coverage than the Bonus

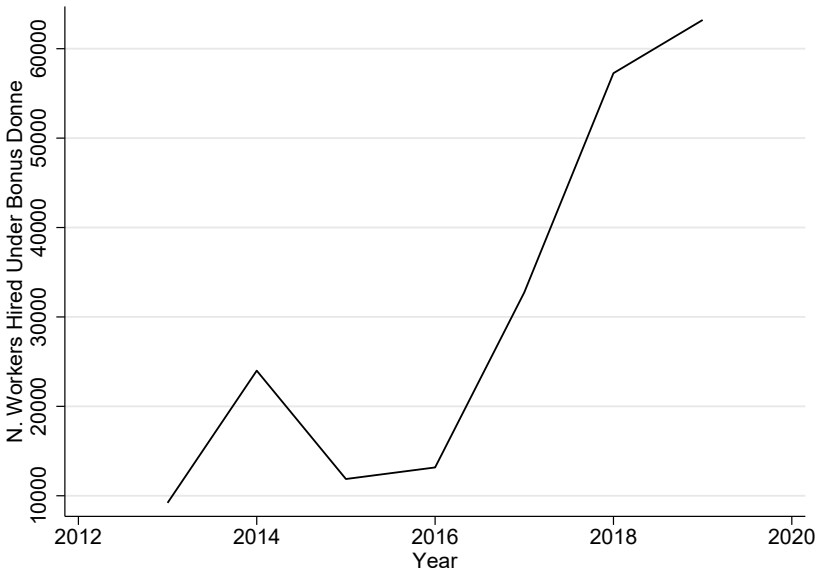
[36]The reform introduced an increase in retirement age and a transition from a defined benefit to a defined contribution scheme for older cohorts.

[37]The law decreased the maximum number of years during which a worker could be employed at a firm under a fixed-term contract. It also increased the minimum time span before a fixed-term worker could be re-hired by the same employer under a fixed-term contract.

Donne. Figure A.2 Panel (a) shows the evolution of Google searches for the item “Bonus Donne” over time. Google searches were very few in 2013 (when the subsidy came into force) and remained low through the first months of 2014. Searches eventually increased and reached a peak in 2015, then remained relatively stable at lower levels from 2016 onwards. Public interest towards this hiring subsidy was significantly lower than for the other, more controversial policies discussed above, as illustrated in Panels (b) and (c) of Figure A.2. Panel (b) compares Google searches for the item “Bonus Donne” (pink line) to searches for “Pensioni Fornero”, the pension reform (blue line). During the announcement and parliamentary discussion of the pension reform in 2012, Google searches for this item reached their maximum, whereas searches for ”Bonus Donne” were very small. This gap is even more striking when Google searches of ”Bonus Donne” are juxtaposed with those for “Article 18” in Panel (c). Such low coverage—hence public knowledge—of the subsidy may be among the reasons for its low take-up, as discussed in Section 3.

A.2 Appendix Figures

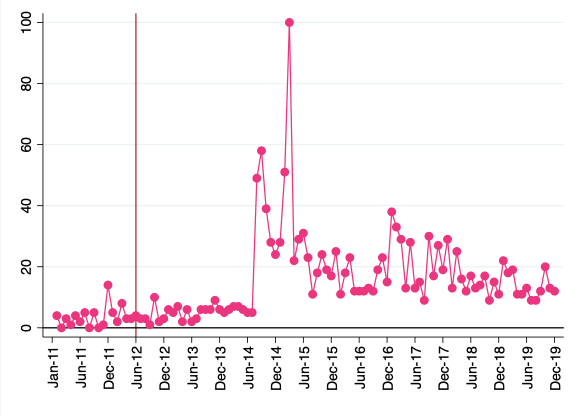
Figure A.1: Subsidy Take-Up Over Time



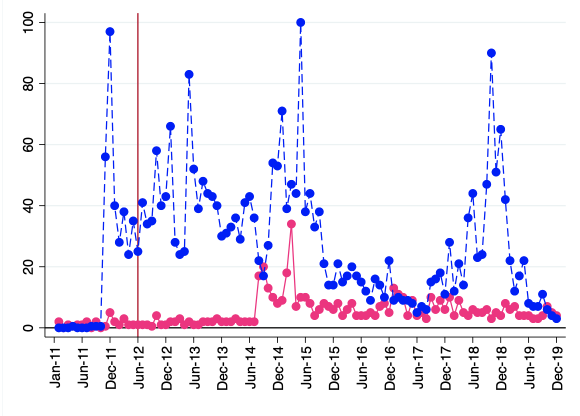
Notes: Total number of workers hired every year with the Bonus Donne subsidy, 2013-2019. The plot also includes men older than 50, for which the same payroll tax cut applied.

Figure A.2: Google Searches, Bonus Donne Versus Other Items

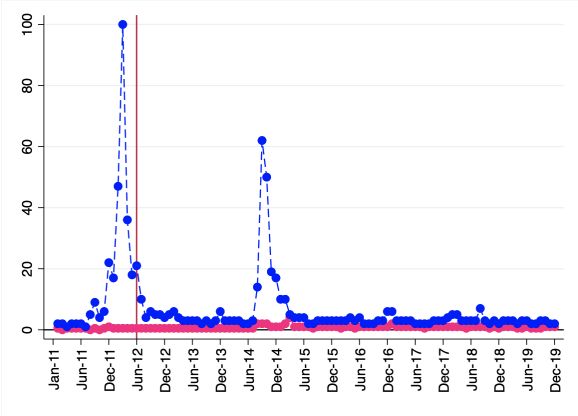
(a) Google Searches for Item "Bonus Donne"



(b) Google Searches for Item "Bonus Donne" (pink) and "Pensioni Fornero" (blue)



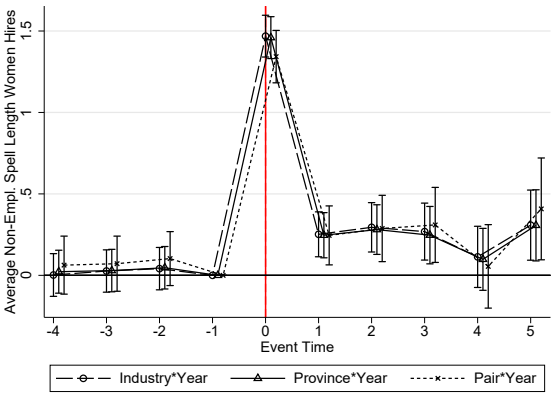
(c) Google Searches for Item "Bonus Donne" (pink) and "Articolo 18" (blue)



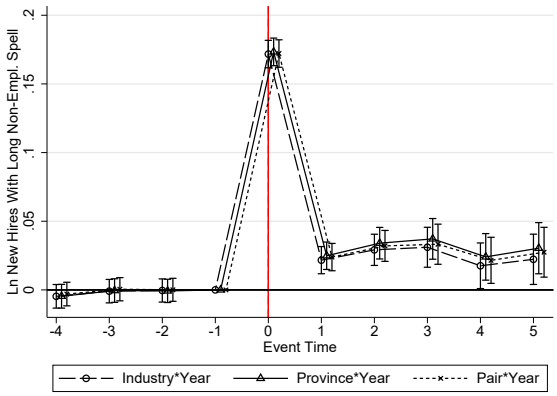
Notes: Google searches between January 2011 and December 2019, sourced from Google Trends. The vertical axis reports the monthly number of Google searches relative to the highest point in the period, normalized to 100. The vertical line denotes the month of approval of the Bonus Donne. The pink time series denotes searches for item "Bonus Donne". The blue time series denotes searches for items "Pensioni Fornero" (referring to the pension reform included in the same law package as the Bonus Donne) in Panel (b) and "Articolo 18" (referring to concurrent labor market reforms) in Panel (c).

Figure A.3: Robustness: Fixed Effects

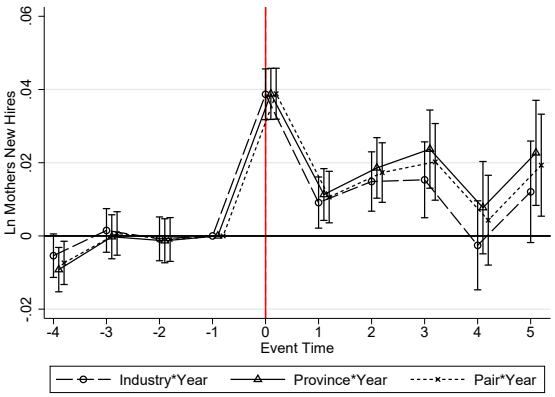
(a) Previous Non-Employment Period



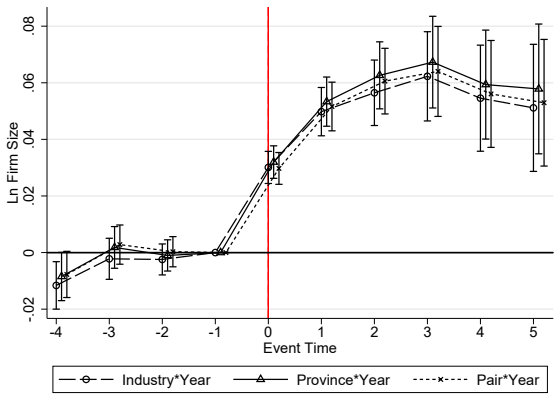
(b) Female Hires with Long Non-Employment Spells



(c) Mothers



(d) Total Firm Employees

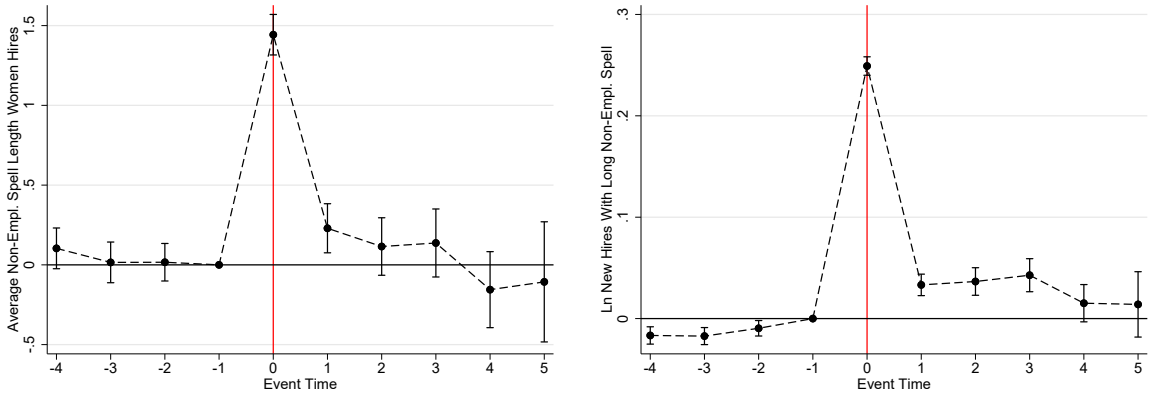


Notes: The figure plots coefficient estimates from Equation 2 denoting differential effects for treated firms using the Bonus Donne versus matched control firms, before and after subsidy adoption (event time = 0). In each specification, the model is augmented with either 2-digit industry*year fixed effects (long-dashed lines), province*year fixed effects (solid lines) or matched treated-control firm pair*year fixed effects (short-dashed lines). Panel (a): The outcome is the average length of non-employment (in years) for new female hires; Panel (b): The outcome is the natural logarithm of (1+) the number of female new hires from long-term (≥ 3 years) non-employment; Panel (c): The outcome is the natural logarithm of (1+) the number of newly hired mothers at the firm (mothers defined as workers that had maternity leave in the previous four years); Panel (d): The outcome is the natural logarithm of the total number of employees at the firm. The red vertical line denotes the period when treated firms hire at least one worker through the Bonus Donne for the first time. The solid vertical lines denote 95% confidence intervals. Standard errors clustered at the firm level. The underlying event study regression control for firm, year and time from-event fixed-effects. Standard errors clustered at the firm level. Matching between treated and control firms is implemented using propensity score matching based on firm-level characteristics measured in the three years preceding the treated firm's first use of the subsidy: average log weekly wage, log total employment, female employment share, and quartiles of the number of hires. In addition, control firms are required to hire at least one female worker in the same calendar year in which their matched treated firm first hires under the Bonus Donne (see Section 5 for details)

Figure A.4: Robustness: Within-Treated Analysis

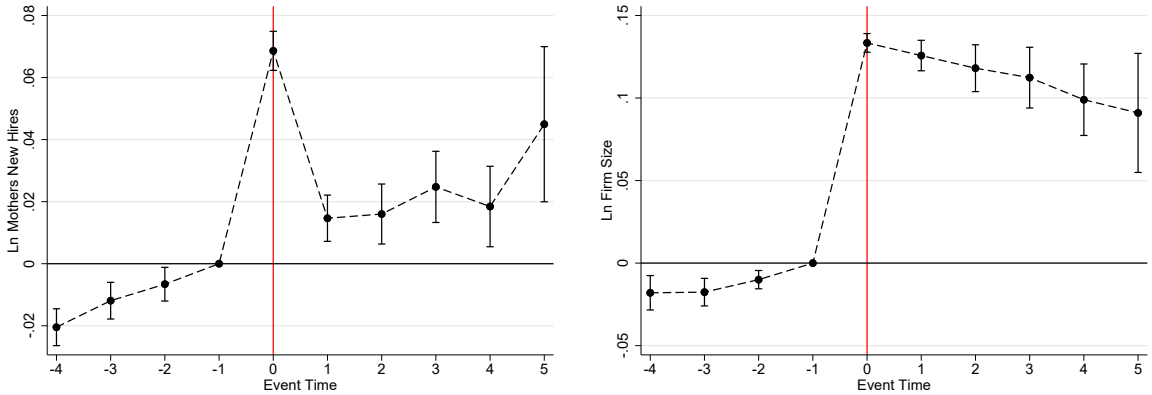
(a) Previous Non-Employment Period

(b) Female Hires with Long Non-Employment Spells



(c) Mothers

(d) Total Firm Employees



Notes: Robustness analysis exploiting staggered treatment timing within treated firms. Event studies estimated using the *eventstudyinteract* command in Stata (Sun, 2021). Coefficient estimates from Equation 2 denoting differential effects for treated firms using the Bonus Donne, before and after subsidy adoption (event time = 0). The estimation sample only considers subsidized firms and uses the last treatment cohort (2019) as control group. Panel (a): The outcome is the average length of non-employment (in years) for new female hires; Panel (b): The outcome is the natural logarithm of (1+) the number of female new hires from long-term ($\geq$ 3 years) non-employment; Panel (c): The outcome is the natural logarithm of (1+) the number of newly hired mothers at the firm (mothers defined as workers that had maternity leave in the previous four years); Panel (d): The outcome is the natural logarithm of the total number of employees at the firm. The red vertical line denotes the period when treated firms hire at least one worker through the Bonus Donne for the first time. The solid vertical lines denote 95% confidence intervals. Standard errors clustered at the firm level.

A.3 Appendix Tables

Table A.1: Additional Firm Descriptive Statistics

	(1) Bonus Donne Firm	(2) Average Italian Firm
	Panel (a): Firm size distribution	
N. Employees	Firm shares	
1-4	0.31	0.43
5-15	0.44	0.41
16-50	0.18	0.12
51-250	0.06	0.03
≥ 251	0.01	0.00
	Panel (b): Geographic distribution	
Area	Firm shares	
Center	0.21	0.22
Islands	0.14	0.09
North-east	0.14	0.21
North-west	0.17	0.28
South	0.34	0.20
Number of firms	26,497	660,764

Notes: The table shows a breakdown of firms by size bins (Panel a) and macro-area (Panel b), comparing firms that have used the Bonus Donne for at least one female hire (Column 1) to other Italian firms that have never made use of the subsidy (Column 2). Data averaged between 2013 and 2019.

Table A.2: Firm Breakdown by Sector

	(1) Bonus Donne Firm	(2) Average Italian Firm
Mining and Quarrying	0.04	0.14
Manufacturing	19.60	21.15
Construction	4.54	12.21
Wholesale and retail trade; repair of motor vehicles	21.75	19.56
Transportation and storage	2.38	4.11
Accommodation and food service activities	20.84	18.81
Information and communication	2.62	2.24
Financial and insurance activities	1.10	1.30
Real estate activities	0.54	0.52
Professional, scientific and technical activities	4.70	4.71
Administrative and support service activities	7.14	5.24
Education	2.33	1.20
Human health and social work activities	5.86	3.02
Arts, entertainment and recreation	1.15	1.05
Other service activities	5.41	4.73

Notes: The table shows a breakdown of firms by sector, comparing firms that have used the Bonus Donne for at least one female hire (Column 1) to other Italian firms that have never made use of the subsidy (Column 2). Data averaged between 2013 and 2019.

Table A.3: Worker Descriptive Statistics

	(1) Bonus Donne Hire	(2) Average Female Hire
Mean age	38.55	35.19
Mean ln weekly wage last job	5.76	5.75
Mean ln weekly wage	5.76	5.69
Average non-empl. spell (years)	6.08	1.74
Mother	0.07	0.06
Hired with open-ended contract	0.27	0.43
Hired with full-time contract	0.31	0.45
Blue-collar	0.61	0.57
White-collar	0.39	0.36
Number of workers	183,615	10,043,297

Notes: The table compares female workers hired through the Bonus Donne (Column 1) to other female new hires from non-employment not through the Bonus Donne (Column 2) between 2013 and 2019. Ln weekly wage is the natural logarithm of the weekly wage earned by workers at the Bonus Donne firm. Ln weekly wage last job is the natural logarithm of the weekly wage earned by workers in their job prior to the Bonus Donne firm, if any. Mothers defined as workers that had maternity leave in the previous four years.

Table A.4: Good-Match versus Bad-Match Firms, Terciles

	(1) Ln Subsidized Female Hires	(2) Ln New Hires From Long-Term Non-Employment	(3) Ln Mothers New Hires	(4) Ln Total Firm Size
Estimate	0.0291**	0.0208**	0.0180**	0.0108
	(0.0128)	(0.0095)	(0.0082)	(0.0152)
Observations	46,002	46,002	46,002	46,002
R-squared	0.356	0.742	0.664	0.925

Notes: The table reports the estimated γ_4 coefficient from Equation 3, denoting the differential effect between treated firms hiring a more versus less productive worker through the Bonus Donne (good-match versus bad-match firms). Good (bad) match firms are those where the modal Bonus Donne hire is in the first (third) tercile of match productivity. Match productivity computed as the residual of a Mincer wage regression (see Section 6 for details). The coefficient is estimated in the post-treatment periods 1 to 5, thus excluding the treatment take-up period. Column (1): The outcome is the natural logarithm of (1+) the number of female workers hired under the Bonus Donne; Column (2): The outcome is the natural logarithm of (1+) the number of female new hires from long-term ($\geq$ 3 years) non-employment; Column (3): The outcome is the natural logarithm of (1+) the number of newly hired mothers at the firm (mothers defined as workers that had maternity leave in the previous four years); Column (4): The outcome is the natural logarithm of the total number of employees at the firm. Standard errors clustered at the firm level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Good-Match versus Bad-Match Firms, Triple Differences

	(1) Ln Subsidized Female Hires	(2) Ln New Hires From Long-Term Non-Employment	(3) Ln Mothers New Hires	(4) Ln Total Firm Size
Estimate	0.0211**	0.0284***	0.0197**	-0.0089
	(0.0094)	(0.0099)	(0.0080)	(0.0165)
Observations	140,448	140,448	140,448	140,448
R-squared	0.353	0.671	0.591	0.929

Notes: The table reports the estimated γ_4 coefficient from Equation 4, denoting the differential effect between treated firms hiring a more versus less productive worker through the Bonus Donne (good-match versus bad-match firms). Good (bad) match firms are those where the modal Bonus Donne hire has above (below) median match productivity. Match productivity computed as the residual of a Mincer wage regression (see Section 6 for details). The estimation sample includes also the matched control firm for each treated firm. The coefficient is estimated in the post-treatment periods 1 to 5, thus excluding the treatment take-up period. Column (1): The outcome is the natural logarithm of (1+) the number of female workers hired under the Bonus Donne; Column (2): The outcome is the natural logarithm of (1+) the number of female new hires from long-term ($\geq$ 3 years) non-employment; Column (3): The outcome is the natural logarithm of (1+) the number of newly hired mothers at the firm (mothers defined as workers that had maternity leave in the previous four years); Column (4): The outcome is the natural logarithm of the total number of employees at the firm. Standard errors clustered at the firm level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Balancing Table, Worker-Level Analysis

	(1) Treated workers	(2) Control workers
Mean age	37.91	38.17
Hired with open-ended contract	0.23	0.23
Hired with full-time contract	0.37	0.34
Non-employment years	4.01	3.92
Apprentice	0.00	0.00
Blue collar	0.53	0.54
White collar	0.47	0.45
Manager	0.00	0.00
Number of workers	13,336	13,336

Notes: The table shows mean values for treated (Column 1) and control (Column 2) workers in the matched sample. Descriptive statistics are computed in the period of the adoption of the subsidy. Non-employment years is the number of years the worker had been in non-employed before behind hired under the Bonus Donne. Matching between treated and control workers is based on the sample female workers hired from non-employment at treated and control firms, using the matched firm sample described in Section 5 and restricting attention to the first year treated firms adopt the subsidy (between 2013 and 2019). We identify female workers hired through the Bonus Donne at treated firms in the initial year of subsidy uptake. Each of these workers is matched to a newly hired female worker in the same year at matched control firms. Matching is based on a propensity-score logit model for the probability that a worker is hired through the subsidy, controlling for worker age, contract type (fixed-term vs. open-ended, part-time vs. full-time), broad occupational categories (manager, white collar, blue collar), and the length of the preceding non-employment spell (see Section 7 for details).

A.4 Model

In this appendix, we introduce a framework that clarifies the connection between the hiring subsidy, firms' hiring decisions, and the learning process that follows when firms hire from a group they would otherwise avoid. This framework clarifies the following points:

1. The policy can have lasting effects on firms' decisions to hire from the targeted group, if it resolves a "learning trap" by exposing firms to workers from that group, thereby inducing firms to learn about the group-firm match quality.

2. If firms learn from interacting with workers of the targeted group, the effect of the policy on firms' future hiring decisions depends on the "quality" of the worker the firm interacts with. In particular, firms that interview (and potentially hire) "good matches" are more likely to interview and hire workers from the targeted group in the future, relative to firms that interview/hire "bad matches".

Our model considers two periods in which firms can potentially hire workers. The hiring process is as follows: first, the firm chooses whether to interview a candidate (based only on the worker's group membership, which informs the firm's prior beliefs about expected productivity); second, if the firm chooses to conduct a costly interview, it observes an additional signal about that worker's productivity, and decides whether to hire them. The expected cost of interviewing a worker is affected by the policy, since the subsidy covers part of the worker's wage; hence, the policy affects interviewing and hiring decisions.

Through the signal observed at an interview, the firm learns about the worker, but also about the group-firm component of productivity. This learning channel implies that first-period interview decisions can have lasting effects by influencing subsequent decisions to interview and eventually hire workers from that group in the second period.

In this model, we also posit that wages are set at the time of hiring, after the firm observes the signal of the worker's productivity. Within this framework, the wage (and in particular wage residuals) reflects the firm's perceived quality of the match at the time of hire.

A.4.1 Environment

A worker i from group g working in firm j has productivity

$$y_{igj} = \bar{y}_{gj} + \theta_{gj} + \theta_i,$$

where $\bar{y}_{gj}$ is a known baseline fit of a person from group g to the firm j; θ_{gj} represents the unknown group-firm component of productivity; and θ_i represents an unknown idiosyncratic component of individual i. We assume that $\theta_{gj} \sim \mathcal{N}(\mu_{gj}, 1)$ and $\theta_i \sim \mathcal{N}(0, 1)$.[38]

Interview decision. The first step of the hiring process is an interview decision for the firm. At this stage, the firm observes only the worker's group membership g and decides to interview them if the expected value of the interview (and subsequent hiring decisions) exceeds some fixed interviewing cost $\kappa > 0$. We denote the expected value of conducting the interview by V_g, an endogenous object that we will describe further below, which depends on features of the learning environment, such as the firm's prior belief about the groupâfirm component of productivity, μ_{gj}. Firm j chooses to interview the candidate from group g if $V_g \geqslant \kappa$.

Hiring decision. Upon interviewing worker i, the firm observes a signal of the worker's productivity, given by

$$\tilde{y} = y_{igj} + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. After observing $\tilde{y}$, the firm forms a belief about y_{igj}:

$$\hat{y}_{igj} = \mathbb{E}\left[y_{igj}|\tilde{y}\right] = \frac{2}{2+\sigma^2}\tilde{y} + \frac{\sigma^2}{2+\sigma^2}(\bar{y}_{gj}+\mu_{gj}) = (\bar{y}_{gj}+\mu_{gj}) + \frac{2}{2+\sigma^2}\left[\tilde{y} - (\bar{y}_{gj}+\mu_{gj})\right]. \quad (6)$$

The firm's payoff from hiring the worker is $\hat{y}_{igj} - w$, where w represents the worker's wage. We posit that the worker's wage (if hired) is proportional to their expected productivity at the firm, so that $w = \alpha\hat{y}_{igj}$, for some fixed $\alpha \in (0, 1)$. Considering a hiring subsidy that pays a proportion $\beta \in (0, 1)$ of the worker's wage, the firm's surplus from hiring the worker is

[38] We normalize the variances of θ_{gj} and θ_i for simplicity. Allowing for different variances would affect the weights used in Bayesian updating but would not impact our results qualitatively.

$s = [(1-\alpha) + \beta\alpha]\,\hat{y}_{igj}$. The firm hires the worker if the surplus is positive.

We can write the value of interviewing a worker from group g (which is subject to the subsidy) as

$$V_g = \mathbb{E}_g\left[\max\left\{0, ((1-\alpha) + \beta\alpha)\,\hat{y}_{igj}\right\}\right],$$

where the expectation is taken with respect to the distribution of $\hat{y}_{igj}$ in group g, given μ_{gj} and σ^2. Note that the hiring subsidy policy does not affect the firm's decision to hire—since the firm always hires workers with $\hat{y}_{igj} \geqslant 0$—but increases the value of V_g, thereby incentivizing firms to interview workers and, through this channel, increase hiring from group g.

Learning. Apart from learning about individual workers' productivity, the firm also learns about the group-firm component of productivity θ_{gj} from observing $\tilde{y}$. This learning is important in our context as it connects first-period hiring decisions to subsequent hiring behavior toward the same group. Specifically:

$$\hat{\mu}_{gj} = \mathbb{E}\left[\theta_{gj}|\tilde{y}\right] = \frac{1}{2+\sigma^2}(\tilde{y} - \bar{y}_{gj}) + \frac{1+\sigma^2}{2+\sigma^2}\mu_{gj}. \tag{7}$$

The firm's belief about θ_{gj} updates upwards if the "surprise" $(\tilde{y} - \bar{y}_{gj})$ is larger than μ_{gj}, and downwards if this "surprise" is smaller than μ_{gj}.

Second-period hiring. Going into the second period, firms differ in what they have learned about θ_{gj}, the group-firm component of productivity. First, a firm may have had *no learning*, which happens if the firm chose not to interview. In that case, second-period hiring mirrors first-period hiring, absent changes in policy that might affect the expected value of interviewing V_g (for example, a change in the subsidy rate).

Firms that interviewed a candidate in the first period observed a signal $\tilde{y}$, and therefore learned about θ_{gj} via the "surprise" $(\tilde{y} - \bar{y}_{gj})$. This learning affects hiring in the second period by changing the distribution of θ_{gj}, so that

$$\theta_{gj} \sim \mathcal{N}(\hat{\mu}_{gj}, \hat{\sigma}^2),$$

where $\hat{\sigma}^2 < 1$ (the closed form expression of $\hat{\sigma}^2$ follows from standard Bayesian updating). This change in the distribution of θ_{gj} affects firms' decisions to interview and hire in the second period. In particular, firms that receive more favorable signals (i.e., higher realizations of $\tilde{y}$) in the first period are more likely to interview workers from group g in the second period, and as a result more likely to hire from this group, relative to firms that receive less favorable signals.

A.4.2 Empirical Implications

This framework highlights a possible long-term effect of the subsidy policy: it can resolve a *learning trap* that prevents firms from hiring from a given group—in our application, women who have been out of the labor market. This occurs in three stages:

Stage 1. In "period 0," before the hiring subsidy is introduced, we assume the firm's beliefs about θ_{gj} are given by $\theta_{gj} \sim \mathcal{N}(\mu_{gj}, 1)$, with μ_{gj} small enough that

$$V_g = \mathbb{E}_g \left[\max \left\{ 0, (1-\alpha)\hat{y}_{igj} \right\} \right] < \kappa.$$

(To see that such a small μ_{gj} exists, note that the distribution of $\hat{y}_{igj}$ is increasing in μ_{gj}, and V_g is an increasing function of $\hat{y}_{igj}$.) Because the value V_g of interviewing a candidate from group g is so small, such candidates are never interviewed. Consequently, no learning occurs, and, in the absence of any policy, candidates from group g are never hired (even in later periods).

Stage 2. In "period 1," a hiring subsidy policy $\beta > 0$ is introduced, where β is large enough that

$$V_g = \mathbb{E}_g \left[\max \left\{ 0, ((1-\alpha) + \beta\alpha)\, \hat{y}_{igj} \right\} \right] > \kappa.$$

With the introduction of this policy, interviewing a candidate from group g becomes profitable. In that period, the firm interviews a candidate and updates it beliefs about θ_{gj}, as described in the learning environment above.

Stage 3. In "period 2," the subsidy policy is removed, so that $\beta = 0$. However, firms' beliefs about θ_{gj} have been updated due to the signals observed during interviews in the previous

period. Some of these firms received signals that were "bad surprises," leading them to continue t interviewing–and therefore not hiring–workers from group g, in the absence of the subsidy policy. Other firms observe sufficiently positive surprises, leading to higher values of $\hat{\mu}_{gj}$ and, consequently, to interviewing and potentially hiring candidates from group g in period 2. On average, the policy can generate persistent increases in hiring of workers in group g, even after the subsidy has been removed. While we document persistent changes in firms' hiring behavior following their use of the subsidy, we do not directly observe the removal of the policy in our empirical setting; this stage therefore illustrates the mechanism through which temporary subsidies can generate such persistence.

These predictions are consistent with our empirical findings in Section 5, where we document persistent increases in firm-level hiring of women with extended labor market interruptions following firms' initial use of the subsidy.

"Good matches" versus "bad matches" in firm learning. An immediate implication of Equation (7), which describes the learning that follows after the firm interviews a worker in group g, is that firms that observe more positive "surprises" $(\tilde{y} - \bar{y}_{gj})$ form higher beliefs $\hat{\mu}_{gj}$, and are consequently more likely to hire workers from group g in the next round of hiring, via an increased propensity to interview candidates from this group.

We test this empirical prediction of the learning model in Section 6, where we define "good matches" / "bad matches" as workers with high / low realizations of the surprise term. Note that the value of the surprise term $(\tilde{y} - \bar{y}_{gj})$ is positively related to the estimate $\hat{y}_{igj}$, which determines the wage offered by the firm. In our empirical analysis, we therefore proxy this surprise term using the Mincerian wage residual of the worker at the time of hiring.