DO NON-COMPETE AGREEMENTS HELP OR HURT WORKERS? EVIDENCE FROM THE NLSY97*

Bhargav Gopal †

Xiangru Li

Luke Rawling

January 21, 2026

Abstract

While non-compete agreements (NCs) are common, the incentives driving their use across various segments of the labor market and effects on workers remain poorly understood. We show theoretically that for non-liquidity constrained workers, NCs compress the external wage profile and encourage firm-provided industry-specific human capital investments, which raises within-job wage growth. This model predicts NCs are used among workers with highly transferable skills. We also present a liquidity constraints model to understand NC usage among lower-wage workers. This model shows that NCs reduce costly turnover, but do not incentivize human capital investments. Lower-wage workers who sign NCs receive a compensating wage differential, but contracted wages will fall below marginal product, lowering the labor share of income. Using novel data on NC usage from the NLSY97, we find NCs are more commonly used among higher skilled workers who perform more sophisticated tasks and have higher returns on experience, consistent with NCs being used to resolve hold-up problems. Using a stacked difference-in-differences research design, we find that NCs raise job tenures by 6% and immediately increase wages by 10%. Six years after signing, the wage premium falls to 5%. NCs increase wage growth for higher-wage workers but reduce wage growth for lower-wage workers. We find no evidence of increased investment in formal training or upgrading of skills, suggesting investments prompted by NCs are likely informal. Our findings caution against blanket bans on NC usage.

*We are grateful to Sandra Black, Bentley MacLeod, Suresh Naidu, and Bernard Salanié for their guidance. We also thank Jean-Etienne de Bettignies, Robert Clark, Evan Starr, and Veikko Thiele for thoughtful comments. We thank Odeya Leikin for excellent research assistance. Gopal acknowledges funding from the Social Sciences and Humanities Research Council of Canada Grant Number 430-2024-00035. Rawling acknowledges funding from the Social Sciences and Humanities Research Council of Canada (Canada Graduate Scholarship – Doctoral).

†Corresponding Author. Email: b.gopal@queensu.ca. Mailing Address: 143 Union St #441, Kingston, ON K7L 3N6. Phone: 8184513475

1 Introduction

Nearly 15% of U.S. workers are bound by non-competition agreements (NCs), and they are used across various segments of the labor market. Despite their commonality – NC usage is comparable to unionization prevalence in the U.S. – there is limited theoretical understanding of why workers voluntary sign NCs despite their restrictive nature. In this paper, we develop a class of fixed-wage contract choice models to understand the incentives driving NC usage among lower-wage workers, higher-wage workers, and executives.

The primary insight is that NCs present a tradeoff between investment incentives and allocative efficiency (e.g., Shi 2023; Gopal, Li, and Rawling 2026). In our model, NCs prevent transitions to industry-competitors, which incentivizes firms to make industry-specific human capital investments in workers. NCs function as a labor market institution that compresses the external wage profile, as in Acemoglu and Pischke (1999), which encourages firms to provide transferable skills. However, NCs cause allocative inefficiencies by restricting job mobility to industry-competitors.[1]

For workers who are not liquidity constrained, our model predicts NCs are used when skills are highly portable, and for signers, faster wage growth due to human capital appreciation and longer job tenures. In the flavor of Becker (1962), workers pay for industry-specific training by taking wages below marginal product initially and then receive marginal product when trained, so wages rise with tenure. In our setting, wages equal marginal product because investment is non-contractible so committing to deferred benefits encourages the firm to invest. NCs therefore reduce turnover, incentivize non-contractible human capital investments, and generate a larger tenure premium, as empirically observed in Lavetti, Simon, and White (2020).

For liquidity-constrained workers who are not willing to accept deferred benefits, NCs function to raise profits by reducing costly turnover. While these workers may initially receive a compensating wage differential for accepting restricted future mobility, wages will eventually fall below marginal product, resulting in lower wage growth, wage markdowns, and a lower labor share of income. If liquidity constrained workers also have lower returns on human capital investments, then NCs will not encourage firm-provided investment. We view this model as applicable to understanding NC usage in low-wage markets.

In contrast, for executives where NCs can be renegotiated and bought out with tailored legal advise, NCs achieve efficient investments and mobility when firms have full bargaining power. This theory may explain why approximately 60% of executives have NCs (e.g., Kini, Williams, and Yin 2021). Overall, while NCs may be pairwise-efficient in various segments of the labor

[1] NCs do not completely deter job mobility, as in Gottfries and Jarosch (2023), because we allow for movement to non-industry competitors. Nearly two-thirds of job transitions are between industries for the workers aged 32- 38 in our sample, so this choice is empirically grounded.

market – contributing to their broad usage – they have varied distributional implications. These nuanced effects of NCs caution against a blanket ban on the enforcement of NCs, as recently proposed by the US Federal Trade Commission and supported by academic research (e.g., Shi 2023).

To test the predictions of our contract choice model, we draw on novel data on NC usage from the National Longitudinal Survey of Youth (NLSY97). The NLSY97 is a nationally representative longitudinal survey that tracks a cohort of individuals who were teenagers in 1997, and runs every two years starting in 2011. By 2017, when the survey first includes a question on NC status, the sampled individuals are between the ages of 32 and 38 — an ideal period for studying labor market outcomes as respondents are in their prime working years. More than 90% are "Very Confident" in the accuracy of the response, supportive of the assumption that workers are aware of contract terms in our theoretical model.

Several facts support the theory that NCs are typically used when skills are more transferable. First, as in Rothstein and Starr (2022), we observe that NC signers tend to be college-educated and higher-wage workers, who are less likely to be liquidity constrained. Second, we find that signers perform more sophisticated tasks such as problem solving, mathematics, and management. Third, linking with O*NET data, we see that NC usage is more common in occupations requiring non-routine cognitive skills, which is associated with skill transferability (Gathmann and Schönberg 2010). Fourth, we see that NCs are used in occupations and industries with higher returns on experience. As in Topel (1991), we think of the returns on experience as the returns on general skills, and we estimate these returns using the same two-step procedure.

We use a stacked difference-in-differences event study to study the dynamic effects of signing NCs on wages, job mobility, and employer-provided investment (e.g., Gormley and Matsa 2011; Cengiz et al. 2019; Jeffers 2023; Johnson, Lavetti, and Lipsitz 2025). We define a cohort c as all workers moving to a new job with known NC status at time c. Within cohort c, the treated group consists of workers who report signing an NC for the first time in year c and the control consists of workers who never report signing an NC over the sample period.[2] This approach creates a panel dataset for each cohort $c \in \{2015, 2017, 2019, 2021\}$, each with its own treatment and control group, which we then "stack" together. This approach addresses the challenges that signing an NC is often coincident with job mobility, which itself is linked to wage increases (e.g., Topel and Ward

[2]Since the NLSY97 only begins tracking NC status in 2017, we assume that if an individual reports signing an NC in 2017, they had it from the beginning of their job tenure. This assumption allows us to estimate the medium-term effects of signing an NC, even though the dataset does not capture NCs from the job's start date. As a sensitivity check, we re-estimate our results using only the 2017, 2019, and 2021 cohorts—workers who started new jobs in those survey waves. For these workers, we directly observe their NC status at the time of hiring, eliminating the need for any assumptions about when the NC was signed. The results from this restricted sample closely align with our main findings, reinforcing the validity of our longer-term estimates. If an individual signs multiple NCs over the sample period, the individual is only included in the treated cohort corresponding to the earliest recorded use of the agreement.

1992), that NC signers have characteristics associated with higher wages (e.g., Starr, Prescott, and Bishara 2021; Rothstein and Starr 2022), and the issues that arise in two-way fixed effects models with staggered treatment timing (e.g., Callaway and Sant'Anna 2021).

Our stacked event-study estimates indicate that signing an NC raises wages by 9.4% within one year, a magnitude comparable to the returns to an additional year of education (e.g., Card 1999). This wage premium persists for at least six years, though we observe that it declines by approximately 1% per year, indicating NCs lower wage growth. We also observe that NCs lower job mobility, as predicted by our theoretical model. On average, NCs increase job tenure by 6% and reduce job mobility by 12%, aligning with the point estimates reported in Lavetti, Simon, and White (2020). Despite theoretical predictions that NCs encourage firm-provided training, we find no significant effects on formal measures of employer-provided training. Similarly, we find no detectable effects on job dissatisfaction, working hours or systematic upgrading of job tasks at conventional levels. We interpret the persistently higher wages earned by NC signers as reflecting higher productivity from informal training or on-the-job learning (e.g. Mincer 1962).

We are confident in the validity of our causal estimates, as our event-study design reveals no pre-trends for wages, job mobility, or other key outcomes. Our conclusions are robust to including a rich set of controls, such as firm size, industry, and occupation, to account for alternative explanations such as NC signers systematically sorting into higher-wage jobs.[3] We find no evidence that NC signers are immediately assigned more sophisticated tasks upon signing, suggesting unobserved shocks to worker productivity are not a potential omitted variable. To further address endogenous selection into an NC, we use "later-treated" workers as a control group, showing our results are robust when comparing individuals who sign NCs in a given year to those who will sign one in the future. Using longitudinal data from 2017 - 2021, we find an NC wage premium from cross-sectional wage regressions and a difference-in-differences switchers design which leverages within-individual variation into and out of NCs, though the premium declines in this specification.

We find heterogeneity in the effects of NCs across various dimensions, including education, income, and race. For college-educated, higher-wage, and White workers, signing an NC is associated with a large and persistent wage premium, reaching as high as 12.4% for higher-wage workers. This dynamic aligns with our model's prediction that non-liquidity constrained workers receive compensation for human capital investments induced by NCs. For lower-wage, non-college-educated, minority, and part-time workers, an initial wage premium is followed by lower

[3]The NLSY97 does not collect detailed information about firm-characteristics, preventing us from explicitly studying whether NC signers join higher productivity firms. However, we note that in the cross-section, there are no statistically significant differences in firm size – a variable highly correlated with firm productivity in labor market models (Manning 2013). The NLSY97 also does not collect information on non-disclosure agreements, which are often bundled with NCs (Balasubramanian, Starr, and Yamaguchi 2024). To the extent that this bundling occurs within our data, our estimates should be interpreted as identifying the causal effects of this broader set of restrictions, rather than NCs in isolation.

wage growth over time. This pattern is suggestive of fewer human capital investments and where the initial premium functions as a compensating differential for accepting restricted job mobility.

Relationship to Literature: Our model closely relates to Shi (2023) and Gopal, Li, and Rawling (2026), who also show that NCs present a tradeoff between industry-specific investment incentives and allocative efficiency. While Gopal, Li, and Rawling (2026) consider an environment where outside options are public information and wages are determined through bargaining, this paper studies fixed-wage contracts and focuses on settings where workers have asymmetric information on outside options. Here firms under-invest because trained workers can quit and other firms can use these portable skills, and not because workers command a higher wage after training.

Our model differs from Shi (2023) in several meaningful ways. First, our model features two types of entrants to capture the realistic feature that NCs only prevent transitions to industry-competitors, while still allowing movement to non-industry competitors. This choice makes industry-specific skills de-facto firm specific, as in Acemoglu and Pischke (1999), and presents NCs in a less restrictive light than Gottfries and Jarosch (2023) where NCs deter all job mobility. Second, in Shi (2023) entrants have private information on match quality which prevents fully state-contingent buyout terms and rules out efficient ex-post renegotiation. As in Posner, Triantis, and Triantis (2004), in our renegotiation model where outside options are public information and buyout payments are feasible, NCs achieve the social planner's solution when firms have full bargaining power, providing a potential explanation for the prevalence of NCs in executive contracts.[4] Third, we model liquidity constrained workers who are unwilling to pay for industry-specific skills by taking wages below marginal product. As in Johnson and Lipsitz (2022), where wages are downward constrained due to institutional features such as minimum wages, we show that NCs will be used to raise profits even absent human capital investments.[5] This theory provides a potential explanation for NC usage in low-wage labor markets.

Beyond the immediate context of NCs, a large literature examines how contracts affect employer incentives to provide general skills. In this regard, the most closely related paper is Acemoglu and Pischke (1999), who show that institutional features of the labor market – such as

[4]Within the class of renegotiation models, we show that when buyout fees are constrained to be common legal remedies such as expectation damages or training repayment programs, NCs result in over-investment and under-separation relative to socially optimal.

[5]See also Gottfries and Jarosch (2023) and Potter, Hobijn, and Kurmann (2024) who embed NCs into Burdett-Mortensen–style models of wage posting and show theoretically how NCs can reduce wages. Johnson, Lavetti, and Lipsitz (2025) show theoretically that the wages of those bound by enforceable NCs is ambiguous as there are two counterveiling forces: faster human capital appreciation creates upward pressure while reduced job mobility creates downward pressure. Meccheri (2009) shows that NCs increase employer incentives to provide general training in an environment where turnover is never efficient. See Rubin and Shedd (1981) for a Becker (1962) style model where NCs encourage firms to provide general investments when workers are liquidity constrained and cannot finance general investments by taking wages below marginal product. In a context without investments, Guimaraes, Pessoa, and Ponczek (2025) present a model in which non-compete agreements with transfer fees can lead to allocative inefficiencies.

adverse selection among job seekers – compress the external wage profile and encourages employer investment in general skills. We present NCs as a labor market institution that transforms industry-specific skills into specific skills, creating similar incentives for employer investment. Unlike Becker (1962) and Acemoglu and Pischke (1999) where employers offer contracts consisting of wage-training packages, we model employer human capital investments as non-contractible (e.g., Che and Hausch 1999). For non-liquidity constrained workers, deferred compensation arises in equilibrium to incentivize non-contractible investments (e.g., Lazear 1979). Our choice to study how fixed-wage employment contracts influence investment incentives follows the spirit of MacLeod and Malcomson (1993), and our decision to endow the worker with asymmetric information about the outside option closely mirrors Hashimoto (1981). As in Hashimoto (1981), transaction costs prevent the worker from credibly revealing his outside option to the employer, preventing renegotiation and resulting in inefficient turnover.[6]

Our empirical analysis contributes to the literature by using panel data to understand how signing NCs affect medium run outcomes. Prior work has focused on the effects of NC regulation rather than the direct worker-level effects of signing NCs (e.g., Johnson, Lavetti, and Lipsitz 2025; Lipsitz and Starr 2022; Jeffers 2023; Kini, Williams, and Yin 2021; Starr 2019). More limited research leveraging micro-data on the usage of NCs has emphasized the descriptive relationships between signing NCs and various worker characteristics (Shi 2023; Starr, Prescott, and Bishara 2021; Rothstein and Starr 2022; Balasubramanian, Starr, and Yamaguchi 2024). Leveraging the panel structure of the NLSY97 pushes the frontier by allowing us to use individual fixed effects to account for time-invariant unobserved heterogeneity and leverage within-worker changes in NC status, improving causal identification. It also enables us to examine how NCs affect wage growth, rather than cross-sectional wage differences, and allows us to track the effects of signing an NC even if a worker later changes jobs.

Our empirical analysis is most closely related to Lavetti, Simon, and White (2020), who use micro-data on NCs among physicians and find that NCs increase earnings by 35% after 10 years of tenure. Leveraging the panel structure of their data to run a Topel-style decomposition, they find that NCs raise the return to tenure but not experience. Using a similar decomposition, we find that NC usage is more common in occupations with high returns on experience, and that the wage gains from signing NCs are concentrated among job stayers. They find NCs increase job duration by 12%, which align with our point estimates as well.[7] Similarly, Lavetti, Simon, and White

[6] As stated on Page 3, "Alternatively, the worker may perceive that the value of his leisure increased in the second period. If the worker could convince the employer of this increase, the employer might increase his wage to retain the worker. Given the absence of readily available measures of the value of leisure, such a transaction may be too costly to occur."

[7] These mobility point estimates are substantially smaller than those reported in contemporaneous work by Cowgill, Freiberg, and Starr (2024), who run a large field experiment with two finance firms and find that removing an NC increases mobility between competing employers by 30-57%.

(2020) find that NCs reduce hold-up by incentivizing owners to increase referrals to physicians; our results complement this finding by showing NCs encourage firms to provide informal training.

Our finding that NCs raise earnings using cross-sectional and panel data techniques has many parallels to the more mature literature on the union wage premium, as recently surveyed by Jäger, Naidu, and Schoefer (2025). Cross-sectional Mincerian regressions typically estimate a union wage premium of roughly 15–20 percent in the U.S. private sector during the late twentieth century (e.g., Gregg Lewis 1986; Freeman and Medoff 1985). In contrast, longitudinal switcher and worker fixed-effects designs yield smaller but still positive estimates on the order of 5–10 percent, consistent with positive selection into union jobs (e.g., Freeman 1984; Lemieux 1998). Similarly, using data from 2017-2021, we find the NC wage premium to be 14 - 20 percent using Mincerian controls and slightly above 5% using the switchers design. Our preferred stacked difference-in-differences specification using data from 2013 - 2021 yields an NC premium of approximately 10%, which matches the cross-sectional wage premium when augmenting Mincerian controls with industry and occupation fixed effects.

The paper proceeds as follows. Section 2 lays out the theoretical framework. Section 3 shows descriptive statistics and the empirical specification. Section 4 presents the results. Section 5 concludes.

2 Theoretical Framework

2.1 The Model

We study a two-period model featuring a risk-neutral firm F and a risk-neutral worker W, whose timeline is described in Figure I. The worker discounts future compensation by $\beta \leq 1$, while the firm values current and future profits equally. At the beginning of period 1, the firm and worker agree to a contract consisting of an NC $\delta \in \{0,1\}$, a fixed wage w_δ, and an up-front transfer B_δ from the firm to the worker. Production in the initial period is normalized to zero and we allow the worker to earn initial period wages below marginal product, or $B_\delta < 0$. The NC restricts the worker, upon separation in period 2, from joining firms in the same industry as F. We initially assume that the contract is rigid: the wage and the NC cannot be renegotiated in period 2 because of the worker's asymmetric information on his outside option. This assumption prevents the firm from tailoring the employment contract to realized conditions and thus results in inefficient turnover.

Because investment is non-contractible, once the contract is signed the firm chooses $i_\delta \geq 0$ to maximize its expected profits, incurring a cost of $\frac{1}{2}i_\delta^2$. An additional unit of investment raises the worker's productivity within the firm by r while raising the worker's productivity at industry competitors by ρ. An NC prevents the worker from moving to industry competitors; as a result,

additional investment by the firm does not raise the worker's outside option when $\delta = 1$.

The firm and worker know the values of r and ρ ex-ante and understand the probability distribution of future outside offers v, but do not observe the realization of v when making the contract or choosing investment. In particular, $v \sim \text{Exponential}(\lambda)$ is drawn at the beginning of period 2 and is privately observed by the worker. We interpret v as the demand for the worker from competing firms, which is private information to the worker, as in Hashimoto (1981). We assume third party firms make an offer equal to their valuation, so third party firms earn zero profits and all the surplus from the employment relationship is captured either by the incumbent firm or the worker.[8] If unrestricted by an NC, the worker's outside option is $v + \rho i_0$; if bound by an NC, his option remains v. As a result, the NC eliminates the worker's access to the within-industry wage premium, but still allows for job mobility across industries.

The worker then decides whether to stay at the firm and receive wage w_δ, or to leave and accept the outside offer. Simultaneously, the firm decides whether to retain or fire the worker. If the firm fires the worker, the firm can hire an untrained worker whose marginal product is zero; as a result, the firm will retain the incumbent worker so long as final period wages are less than or equal to marginal product, or $w_\delta \leq r i_\delta$. We remain agnostic about labor market conditions at the time of contracting. The worker has a reservation utility of μ^0 in the initial period that is commonly known to all parties, and will only accept the contract if expected utility exceeds this value. Symmetrically, the firm will only offer a contract if its expected profit is non-negative, which in turn requires the total joint surplus of the relationship to be at least as large as the worker's reservation utility. Throughout our analysis, we assume this condition holds, as no mutually beneficial contract could be formed otherwise. Since the firm acts first, it will choose a contract so that the worker's participation constraint binds. We highlight that although the model features a single firm, it can represent labor markets with varying degrees of competition ex-ante. For example, in a perfectly competitive labor market, μ^0 should be interpreted as the endogenous utility level which incentivizes firms to offer contracts that yield zero expected profit.

We allow both r and ρ to vary freely. When $\rho = 0$, investment is purely firm-specific; when $\rho = r$, it is fully transferable across firms in the industry. In practice, different sectors exhibit different returns to training, and the external value of a worker's training may vary substantially. Importantly, when $\rho > 0$, the firm's investment benefits not only the firm itself but also the worker by improving his outside option, so the investment is *industry-specific* in nature (e.g., Parent 2000). We focus on cases where $r > \rho$, meaning the internal return on investment exceeds the external return. However, the model can also accommodate situations in which a firm's investment increases the worker's productivity more for third parties than for the incumbent firm, as in Pakes and Nitzan

[8] This assumption can be justified in an environment where there are multiple third-party firms within and outside the industry, as in Spier and Whinston (1995).

(1983). We view this case as particularly applicable to highly skilled scientific personnel, where the value of a discovery made at the original firm may be even greater when developed in a new enterprise or spun out into a separate company.

In standard hold-up models, firms underinvest because they anticipate having to share the additional surplus with workers through a higher wage (e.g., Becker 1962; Acemoglu and Pischke 1999). Here, instead, the firm underinvests without an NC due to two reasons. First, the firm does not internalize the benefits to the worker when he leaves for a competitor. Second, the firm's investment raises the worker's outside option and thus the probability of a quit. As a result, the firm captures only a fraction of the marginal social returns to investment and invests sub-optimally.

NCs mitigate this distortion by decoupling investment from the worker's outside option. When $\delta = 1$, the worker's outside option is fixed at v, so increased investment does not raise the probability the worker quits. Furthermore, for any given wage, a worker is less likely to quit with an NC than without an NC. Both of these forces raise the firm's incentives to invest in industry-specific skills. However, NCs also induce inefficient stays or "job lock." When $v + \rho i_1 > r i_1$ and $w_1 > v$, separation to an industry competitor is socially efficient but the worker is contractually barred from doing so. In these instances, the firm's private returns from investment exceed social returns. As a result, if investment is highly specific (large $r - \rho$), an NC may lead to over-investment relative to the socially optimal level.

2.2 The Social Planner's Solution

To establish an efficiency benchmark, we first characterize the allocation chosen by a social planner who controls both the investment and separation decisions to maximize total social surplus. The planner's problem unfolds in two stages. First, the planner chooses an investment level $i \geq 0$ at a social cost of $\frac{1}{2}i^2$. Second, after the outside option $v \sim \text{Exponential}(\lambda)$ is realized, the planner dictates whether the worker stays with the firm or separates. The planner makes this ex-post decision by comparing the total surplus generated in each state. If the worker stays, the surplus is ri. If the worker separates, the surplus is $v + \rho i$.

It is therefore socially efficient for the worker to stay if and only if the surplus from staying exceeds the surplus from separating. We typically assume $r > \rho$, meaning investment is more productive inside the firm.[9] The efficient separation rule is to continue the match whenever

$$ri \geq v + \rho i \quad \iff \quad v \leq (r - \rho)i \tag{1}$$

Letting $\Delta := r - \rho$ denote the productivity gap, the planner retains the match if $v \leq \Delta i$.

[9] When $\rho \geq r$, it is straightforward to show that the planner always separates ex-post and sets $i^\star = \rho$.

Anticipating this efficient separation rule, the planner chooses the investment level i ex-ante to maximize the expected total surplus:

$$\mathscr{S}(i) = \int_0^{\Delta i} (ri)\lambda e^{-\lambda v} dv + \int_{\Delta i}^{\infty} (v + \rho i)\lambda e^{-\lambda v} dv - \frac{1}{2} i^2$$

Using the properties of the exponential distribution, this simplifies to:

$$\mathscr{S}(i) = -\frac{1}{2} i^2 + ri + \frac{e^{-\lambda \Delta i}}{\lambda}$$

The socially optimal level of investment, $i^\star$, is therefore implicitly defined by the first-order condition $\mathscr{S}'(i^\star) = 0$, which equates the social marginal benefit of investment (SMB) with the marginal cost of investment:

$$\underbrace{r - \Delta e^{-\lambda \Delta i^\star}}_{=\text{SMB}(i^\star)} = i^\star \tag{2}$$

Appendix A.1 shows the existence and uniqueness of the investment level as well as the comparative statics of optimal investments and quit probabilities with respect to the parameters of the model.

2.3 Subgame Perfect Equilibrium with an NC

2.3.1 Equilibrium Characterization

The equilibrium is solved by backward induction. Before analyzing the firm's choices, we define the worker's expected utility. In the final period, the worker stays if $v \leq w_1$ and quits if $v > w_1$. The worker's ex-ante expected future utility, $E[U_W(w_1)]$, calculated before the outside option v is realized, is therefore:

$$\begin{aligned} E[U_W(w_1)] &= \int_0^{w_1} w_1 \lambda e^{-\lambda v} dv + \int_{w_1}^{\infty} v \lambda e^{-\lambda v} dv \\ &= w_1(1 - e^{-\lambda w_1}) + e^{-\lambda w_1}(w_1 + 1/\lambda) \\ &= w_1 + \frac{1}{\lambda} e^{-\lambda w_1} \end{aligned}$$

The worker's participation constraint is $B_1 + \beta E[U_W(w_1)] \geq \mu^0$.

For a given wage w_1, the firm chooses investment i_1 to maximize its expected profit, anticipating the worker's separation decision:

$$\max_{i_1 \geq 0} \quad \underbrace{(ri_1 - w_1)}_{\text{Profit if stay}} \underbrace{(1 - e^{-\lambda w_1})}_{\text{Pr(stay)}} - \frac{1}{2} i_1^2$$

The first-order condition yields the firm's optimal investment response function:

$$i_1(w_1) = r(1 - e^{-\lambda w_1}) \tag{3}$$

This expression indicates that in equilibrium, the marginal cost of investment equals the expected private marginal return to investment, which equals the probability the worker stays ($1 - e^{-\lambda w_1}$) multiplied by marginal product r of investment. In the initial period, the firm chooses (w_1, B_1) to maximize its profit. Since the bonus B_1 can be used to transfer utility ex-ante, this problem is equivalent to choosing w_1 to maximize the joint surplus of the match, Σ_1, subject to the constraint that the firm will not fire the worker ($w_1 \leq ri_1$). Although the firm has the option to fire the worker in period 2, this option is never exercised in equilibrium.[10] To see why, observe that if wages are above marginal product in the final period, the firm must earn profits in the first period for the relationship to be profitable. Anticipating that a firing will occur, the firm will make no investments. The worker is well aware of the firm's incentives and thus would be unwilling to earn below marginal product in the first period, so the relationship will not materialize in the first place. We denote this feature as the firm's "viability constraint", which rules out wage-tenure profiles where wages can exceed marginal product, as in the implicit insurance contracts explored in Harris and Holmstrom (1982).

The joint surplus is the sum of the firm's profit and the worker's discounted expected utility:

$$\Sigma_1(w_1) = \left[(ri_1(w_1) - w_1)(1 - e^{-\lambda w_1}) - \frac{1}{2} i_1(w_1)^2 \right] + \beta E[U_W(w_1)]$$

Substituting in the expression for $E[U_W(w_1)]$, maximizing Σ_1 with respect to w_1, and applying the Envelope Theorem yields the following relationship.

Proposition 1 (Optimal Contract Structure)**.** *The optimal wage w_1 and the corresponding investment level $i_1(w_1)$ are related by the equation:*

$$w_1 = ri_1(w_1) + \frac{(\beta - 1)(e^{\lambda w_1} - 1)}{\lambda} \tag{4}$$

For any worker patience $\beta \leq 1$, this contract ensures the equilibrium wage satisfies the firm's viability constraint, $w_1 \leq ri_1(w_1)$, so the firm does not fire the worker.

[10] Only approximately 5% of workers were fired in our dataset in 2017, making this feature of the model relatively realistic.

In Appendix Section A.2, we provide the necessary and sufficient conditions for a non-zero wage. All further analysis imposes $\lambda r^2 > 2$, which provides a sufficient condition. This wage structure creates a back-loaded compensation profile, where workers receive lower compensation early in the job in exchange for higher wages in the future. The steepness of this intertemporal trade-off is governed by the worker's patience β. A more patient worker is more willing to trade present for future earnings, accepting lower initial compensation (B_1) in exchange for a higher future wage (w_1). When $\beta = 1$, the worker earns exactly marginal product in the final period; to satisfy the firm's participation constraint, it must be the case that the worker earns less than marginal product in the initial period ($B_1 < 0$), leading to an upward sloping wage-tenure profile.

2.3.2 Comparison with the Social Optimum

The structure of the private contract deviates from the social optimum because the firm's objective function ignores the external productivity parameter ρ. We formally compare the investment levels when the worker does not discount the future ($\beta = 1$).

Proposition 2 (Private vs. Social Investment)**.** *When $\beta = 1$, private investment with an NC, i_1, is compared to the social optimum, i^*, as follows:*

- *A sufficient condition for under-investment ($i_1 < i^*$) is $1 - \lambda r i^* > 0$.*
- *A sufficient condition for over-investment ($i_1 > i^*$) is $1 - \lambda \Delta i^* < 0$, where $\Delta = r - \rho$.*

The potential for both over- and under-investment stems from how the NC contract misaligns private and social incentives. With a perfectly patient worker ($\beta = 1$), the wage is set to the marginal product, $w_1 = r i_1$. This structure eliminates inefficient quits, but creates two competing distortions. First, the contract creates inefficient stays, which occur when a quit is socially optimal ($v > (r - \rho) i_1$) but does not happen because the worker's outside offer is below the wage ($v \leq r i_1$). In this region, the firm values the marginal return on its investment at r, whereas the planner would have valued it at ρ had the worker moved. Since $r > \rho$, the firm's incentive to invest is stronger than the planner's in this state, creating a force for over-investment. Second, the NC contract misvalues efficient quits. When the worker's outside offer is high enough to induce a quit ($v > r i_1$), the firm's return on its investment is zero. The social planner, in contrast, recognizes that the investment still generates value through the worker's enhanced productivity at a new firm, yielding a social return of ρ. Because the firm completely disregards this post-separation value, its investment incentive is weaker than the planner's in this state, creating a force for under-investment. The ultimate comparison to the efficient level is therefore ambiguous, as it depends on which of these forces dominates. If the risk of quitting to high-value outside offers dominates, the firm's failure to internalize the post-quit value (ρ) can lead to under-powered investment incentives. If, on the

other hand, quit risk is low and investments are highly specific, the firm's desire to reap its full internal return (r) can lead to over-powered investment incentives. The formal proof is provided in Appendix A.3.

2.4 Subgame Perfect Equilibrium Without an NC

2.4.1 Equilibrium Characterization

The equilibrium is characterized by the firm's optimal investment and wage choices. The firm chooses the wage w_0 to maximize the private joint surplus of the match, anticipating its own investment response in the subsequent stage. The joint surplus, Σ_0, is the sum of the firm's expected profit and the worker's discounted expected future utility:

$$\Sigma_0(w_0, i_0) = (ri_0 - w_0)(1 - e^{-\lambda T_0}) - \frac{1}{2}i_0^2 + \beta\left(w_0 + \frac{1}{\lambda}e^{-\lambda T_0}\right)$$

where $T_0 = w_0 - \rho i_0$ is the retention threshold for labor demand conditions v.

First-Order Condition for Investment. The firm chooses i_0 to maximize its own expected profit, $E[\Pi_0] = (ri_0 - w_0)(1 - e^{-\lambda T_0}) - \frac{1}{2}i_0^2$, taking the wage w_0 as given. The first-order condition implicitly defines the firm's investment response function, $i_0(w_0)$:

$$H(i_0, w_0) \equiv \underbrace{r(1 - e^{-\lambda T_0})}_{\text{Marginal Product Effect}} - \underbrace{(ri_0 - w_0)\lambda\rho e^{-\lambda T_0}}_{\text{Hold-up Cost}} - \underbrace{i_0}_{\text{MC}} = 0 \tag{5}$$

This condition characterizes the firm's optimal investment choice, taking the wage w_0 as fixed. The first term, $r(1 - e^{-\lambda T_0})$, is the expected marginal benefit of investment: the gain in the worker's productivity if the match persists multiplied by the probability the match persists. The second term, $(ri_0 - w_0)\lambda\rho e^{-\lambda T_0}$, captures the hold-up problem: by investing more, the firm raises the worker's outside option and thus the likelihood of a quit, making it more probable that the firm will lose its profit margin on the worker, $(ri_0 - w_0)$.

First-Order Condition for Wages. The firm chooses w_0 to maximize the joint surplus $\Sigma_0(w_0, i_0(w_0))$, subject to the viability constraint $w_0 \leq ri_0(w_0)$. The unconstrained optimum, which we denote w^*_{unc}, is the wage that solves the first-order condition $\frac{d\Sigma_0}{dw_0} = 0$, yielding

$$\lambda e^{-\lambda T_0}(ri_0 - w_0) + (\beta - 1)(1 - e^{-\lambda T_0}) + \beta\rho e^{-\lambda T_0}\frac{di_0}{dw_0} = 0 \tag{6}$$

- The first term, $\lambda e^{-\lambda T_0}(ri_0 - w_0)$, reflects the firm's gain from increasing retention: raising the wage increases the probability the worker stays, which is valuable when the worker's marginal product exceeds his wage.

- The second term, $(\beta - 1)(1 - e^{-\lambda T_0})$, captures the cost of transferring surplus to an impatient worker. If the worker is impatient ($\beta < 1$), then increasing the wage reduces joint surplus because the worker undervalues the extra compensation.

- The third term, $\beta\rho e^{-\lambda T_0}\frac{di_0}{dw_0}$, is the indirect benefit of higher wages through investment incentives: a higher wage reduces hold-up risk, encouraging more investment, which in turn raises the worker's future utility by improving his outside option.

Note that this unconstrained solution is only chosen if it is feasible. If w^*_{unc} violates the viability constraint (e.g., if $w^*_{unc} > ri_0(w^*_{unc})$), then the equilibrium wage will be a corner solution. In this case, the firm is constrained and will choose the highest possible wage that still satisfies viability. This is the wage w^*_0 that lies on the boundary of the feasible set, solving the fixed-point equation $w^*_0 = ri_0(w^*_0)$. In Appendix A.4, we show the viability constraint binds for workers who do not discount the future. We also show the existence of the fixed point and the conditions under which it is positive.[11] As in the case with an NC, this optimal contract features a back-loaded compensation profile. The worker's patience, β, determines the steepness of this intertemporal trade-off. A more patient worker is more willing to trade lower initial compensation (B_0) for a higher future wage (w_0) that encourages greater firm investment. This dynamic is most apparent in the limit when $\beta = 1$. In this case, the unconstrained solution would feature a wage above marginal product. Under such an agreement, the firm would be motivated to fire the worker ex-post, so the viability constraint binds and the equilibrium is a corner solution where the worker receives exactly product in the final period. To ensure the firm participates ex-ante, wages are below marginal product in the initial period, again leading to an upward sloping wage-tenure profile.

2.4.2 The Hold-Up Problem: Under-Investment

Proposition 3 (Under-investment without an NC)**.** *In the equilibrium without an NC, the firm's investment level, i^*_0, is strictly less than the socially optimal level, i^*.*

[11] Throughout our analysis, we assume parameter values satisfy both $\lambda r(r-\rho) > 1$ and $\lambda r^2 > 2$. The first condition ensures a positive wage and investment in the equilibrium without an NC, while the second is a sufficient condition ensuring a positive wage in the equilibrium with an NC. These assumptions ensure that we are always comparing two non-trivial contractual outcomes.

The proof establishes this result by showing that under-investment occurs even in the most efficient possible contract without an NC, and this inefficiency is only compounded in more general cases. The best-case scenario occurs when the wage is set at the highest possible level that allows the firm to break even on a retained worker ($w_0 = ri_0$), a situation that arises when $\beta = 1$. This contract is efficient in one dimension: it eliminates inefficient quits by aligning the private separation decision with the social one. Nevertheless, the firm still under-invests because it fails to internalize the full benefits of its investment upon an efficient separation. When the worker leaves, the firm receives nothing, but its investment still generates social value by increasing the worker's productivity at a new firm by ρ. Since the firm ignores this positive externality, its private incentive to invest is strictly lower than the social planner's.

This under-investment problem is compounded when the worker does discount. In this case, the optimal wage is below the worker's marginal product ($w_0^* < ri_0^*$), which re-introduces the problem of inefficiently high quits. This lower wage not only creates excessive turnover but also directly reduces the firm's investment level. Because investment is already suboptimal even in the best-case scenario, it must therefore always be strictly below the social optimum. The formal proof, which establishes this result by analyzing an upper bound for investment, is provided in Appendix A.5.

2.5 Comparing NC to No-NC Contracts: Investment, Wage-Tenure Profiles, Contract Choice

2.5.1 NCs Raise Investment

Proposition 4 (NCs and Investment Incentives)**.** *Let $i_1(w)$ and $i_0(w)$ be the firm's optimal investment response functions to a given wage w with and without an NC, respectively. Let i_1^* and i_0^* be the corresponding equilibrium investment levels.*

(a) For any viable wage w and skill transferability $\rho > 0$, the firm's investment response is strictly higher with an NC: $i_1(w) > i_0(w)$.

(b) When the worker does not discount the future ($\beta = 1$) and the skill is tranferable ($\rho > 0$), the equilibrium investment is strictly higher with an NC: $i_1^* > i_0^*$.

Corollary 1. *For the benchmark case of a perfectly patient worker ($\beta = 1$) and $\rho > 0$, the NC alters the structure of the equilibrium contract in the following ways:*

(a) The equilibrium wage is strictly higher ($w_1 > w_0$).

(b) The equilibrium quit probability is strictly lower ($q_1 < q_0$).

(c) The comparison of the upfront bonuses (B_1 vs. B_0) is ambiguous.

Corollary 2 (Compensating Wage Differential)**.** *If the equilibrium wage with an NC is strictly lower than the equilibrium wage without an NC ($w_1 < w_0$), then the upfront bonus must be strictly higher ($B_1 > B_0$), assuming the worker has a positive discount factor ($\beta > 0$).*

The main proposition establishes that NCs foster higher investment by resolving two distinct disincentives that exist in contracts without this protection. First, for any given level of investment and wages, the baseline probability of a quit is higher without an NC because the worker's outside option is enhanced by their value to competitors. This elevated turnover risk naturally lowers the expected return on any investment the firm makes. Second, the act of investing further undermines the firm's position: raising investment by one unit also raises the probability of a quit. This hold-up effect, where investment actively increases the risk of separation, creates a disincentive to invest. An NC mitigates both problems, providing a stronger private incentive for the firm to invest.

The corollaries explore the consequences of this higher investment. When $\beta = 1$, wages are set to the marginal product ($w_\delta = ri_\delta$). Since an NC leads to higher investment ($i_1^\star > i_0^\star$), it directly follows that it also supports a higher equilibrium wage ($w_1 > w_0$). This higher wage, in turn, unambiguously lowers the quit probability. The effect on the upfront bonus is ambiguous. While the NC contract offers a higher future wage, it also restricts the worker's mobility, reducing the quit option value. The net effect on the worker's ex-ante expected utility is therefore unclear. The second corollary further clarifies this trade-off. It establishes that if an NC contract were to offer a lower wage, it must compensate the worker with a higher upfront bonus, as the worker would face two disadvantages: a lower direct wage and a less valuable option to quit. To offset this double loss, the firm must provide a larger initial payment to satisfy the worker's participation constraint. The formal proofs for the proposition and its corollaries are provided in Appendix A.6.

2.5.2 Incentives to Use NCs

This section explores the conditions under which firms and workers prefer contracts with or without NCs. The choice depends critically on how transferable skills are across firms (ρ) and how much workers value future income (β).

Proposition 5.

- *When skills are entirely firm-specific, the hold-up problem vanishes. The No-NC and NC contracts are identical.*
- *Industry-Specific Skills ($\rho > 0$) with a Patient Worker ($\beta = 1$): The choice of contract is determined by a unique tipping point, $\hat{\rho} \in (0, r)$.*
  - *For all skill generality levels below the tipping point ($\rho \in (0, \hat{\rho}]$), the No-NC contract is strictly preferred.*

  - *For levels above the tipping point ($\rho > \hat{\rho}$), a marginal increase in skill generality decreases the surplus of the No-NC contract, increasing the relative incentive to use an NC contract.*
  - *This tipping point $\hat{\rho}$ is strictly increasing in both the internal return on investment (r) and the quit rate parameter (λ), expanding the range where the No-NC contract is favored.*

- *When the worker is completely impatient ($\beta = 0$), the firm's expected profit is strictly higher under an NC than without an NC. The firm will therefore unambiguously choose the NC contract.*

Corollary 3. *If investment makes a worker more productive for a competitor than for the incumbent firm ($\rho > r$), an NC is always strictly preferred to a No-NC contract.*

This analysis reveals that the choice between the NC and No-NC contracts hinges on a trade-off between fostering investment and allowing for efficient separations. When skills are purely firm-specific ($\rho = 0$), the hold-up problem is absent, rendering the two contracts identical. However, once skills become valuable to competitors ($\rho > 0$), the parties are not indifferent to contract choice.

For a patient worker ($\beta = 1$), skill transferability is a double-edged sword. While a higher ρ increases the value of an efficient quit (a benefit only realized with a No-NC contract), it also exacerbates the hold-up problem, depressing the firm's incentive to invest. Our proposition shows that for low levels of skill transferability ($\rho < \hat{\rho}$), the benefit of enabling efficient quits outweighs the cost of the mild hold-up problem, making the No-NC contract superior.[12] As skills become highly transferable ($\rho > \hat{\rho}$), incentives to use an NC rise. The tipping point $\hat{\rho}$ increases with the internal productivity r and the quit rate parameter λ. A higher r strengthens the firm's internal returns on investment, allowing it to tolerate greater skill transferability before the hold-up costs become too severe, thus expanding the range of ρ where No-NC is preferred. Similarly, a higher λ implies lower expected outside offers (since $E[v] = 1/\lambda$), which reduces baseline quit risks and mitigates the severity of the hold-up problem, again favoring No-NC over a broader parameter space. In the extreme case where investment makes a worker more productive for a competitor ($\rho > r$), an NC is essential for any investment to occur at all, and so will be used.

Worker patience is the other important determinant. When a worker is completely myopic ($\beta = 0$) and values only the upfront bonus, the choice of contract falls entirely to the firm to maximize its own profit. The firm unambiguously chooses the NC because it reduces costly turnover

[12] In results available upon request, we show the No-NC contract is strictly preferred to the spot market – where wages are set after the firm invests. Therefore, the No-NC contract will be used by transitivity.

and protects its investment returns. This result provides a potential explanation for the prevalence of NCs among low-wage or financially constrained workers, who may exhibit high effective discount rates. A large body of research in economics suggests that financial pressures can lead individuals to prioritize immediate cash over future income streams (Lawrance 1991). Our model shows that under such conditions of worker impatience, firms have a clear profit-based incentive to use NCs. The formal proofs are provided in Appendix A.7. In Appendix B, we simulate the model under various parameter values. These simulations visualize our main propositions and corollaries, showing the conditions under which NCs are used, and the effects of NCs on wages, wage growth, quit rates, and investment. Notably, the simulations indicate that when workers discount the future, there are many instances in which NCs reduce wages within the firm (Figure B2 and Figure B3), adding to our theoretical understanding of how NCs can potentially reduce wages (Potter, Hobijn, and Kurmann 2024; Gottfries and Jarosch 2023).

2.6 Extension: Allowing for Renegotiation

In Appendix C, we analyze how equilibrium outcomes change once we allow for renegotiation under different legal remedies. We view this model as a lens to understand the prevalence of NCs in executive contracts (e.g., Kini, Williams, and Yin 2021; Shi 2023), where the buyout payment can be negotiated with tailored legal advice. However, because the NLSY97 contains few data points in executives, we leave this model to the appendix. With unrestricted renegotiation, the firm can adjust wages, buyout payments, or release the worker from the NC after commonly observed demand shocks are realized. In this case, NCs implement the first-best: turnover occurs exactly when efficient, and the firm invests at the socially optimal level, aligning with the theoretical findings in Posner, Triantis, and Triantis (2004). The reason is that the firm can design buyout payments so that it captures the full return on its investment, including the value that accrues to competitors when the worker leaves. By extracting these benefits through the buyout, the firm internalizes all gains from investment, leading to efficient investment and separation decisions.

When legal rules restrict the structure of buyout payments, inefficiencies arise. Under expectation damages, the firm is fully insured against worker turnover, which induces over-investment and too little separation relative to socially optimal. Training Repayment Agreements (TRAPs), where workers are required to repay firm training costs in order to be released from an NC, create similar distortions: if the firm pursues a high-investment strategy, the repayment fee becomes so large that separation rarely occurs, leading to excessive investment. Alternatively, if the firm holds back on investment to keep the repayment fee low enough that separation is sometimes feasible, the resulting investment level is dictated by the firm's private incentives rather than a regime that maximizes total surplus. In either case, investment and turnover are inefficient: the firm either

over-invests to lock in the worker or invests at a level that fails to internalize the full social value of training. Finally, in the absence of an NC, renegotiation yields efficient turnover but persistent under-investment since the firm does not internalize the benefits of industry-specific skills to other employers.

3 Empirical Set-Up

3.1 Data and Descriptive Relationships

We use data from the National Longitudinal Survey of Youth 1997 (NLSY97) to understand the characteristics of NC signers and analyze the effects of such agreements on worker outcomes, including wages, job mobility, and employer-provided investment. This data set is a nationally representative panel that tracks the outcomes of individuals who were ages 12-16 in 1997. The survey runs annually from 1997-2010, and then biannually from 2011-2021. The survey includes information on each individual's employment history, including hourly wages for each job held, as well as detailed information on demographics and job information.

Importantly, the NLSY97 starts measuring whether NCs are used within employment contracts starting in 2017, when survey respondents are between ages 32-36. In 2017, all working respondents are asked whether they currently have an NC. In the following survey years, all individuals who obtain *new* jobs are asked about their NC status. We assume throughout that NC status is fixed for the duration of the employment relationship.[13] We focus on the 2013-2021 time period and individuals who sign NCs starting in 2015, allowing us to estimate the impact of NCs for up to 6 years. We restrict the analysis to employed workers and remove observations with real hourly wages below \$3 or above \$200, following Deming (2017). We construct real hourly wages by deflating nominal hourly pay by annual CPI indices from the Bureau of Labor Statistics, setting 2017 as our base year. When individuals hold multiple jobs in a survey year, we restrict attention to their primary job which we define as the current or most recent employer as of the interview date. If multiple jobs are current, the main job is the one with the longest tenure. We use the survey weights provided by the NLSY to produce estimates that are nationally representative of the target population.

We are interested in the relationships between NCs and various labor market outcomes. We observe the employment history of each worker, including job identifiers and hourly wages at each job, allowing us to assess the impact of signing an NC on wages, wage growth, and job mobility. We consider log wages and job tenure as our main outcome variables throughout the analysis. As

[13] Starr, Prescott, and Bishara (2021) conducts a large survey that asks about the timing of NCs and find that in the vast majority of cases an NC is signed prior to or immediately after starting the job, with only 2.2% associated with promotions or raises.

robustness we explore other measures of job mobility, such as an indicator variable for whether an individual changed main employers between survey years, and come to similar conclusions. The NLSY97 also asks a variety of questions about formal training programs. As a default, we report statistics pertaining to whether an individual was involved in a formal training program, but also consider whether this training was paid for and/or provided by the employer. Full variable definitions are provided in Table E1.

We find that 15 percent of workers in the 2017 cross-section report having an NC in their current job (Table E2). Importantly, over 90 percent of these affirmative respondents report being "Very Confident" in their answer, which supports the modeling assumption that workers are aware of their contract terms and the implications of contract choice.[14] NC usage also varies substantially across the 17 two-digit industries in our sample. Among industries with more than 100 respondents, NCs are most common in Professional and Related Services (28 percent) and least common in Entertainment, Accommodations, and Food Services (7 percent).

In Table I we present summary statistics for the 2017 cross-section, comparing workers with and without NCs. Workers who sign NCs earn higher wages, with an average gap of 21 log points. They also exhibit slightly longer job tenures and are 4 percentage points less likely to change main employers between 2017 and 2019. Since NCs legally restrict within-industry job mobility, we also examine within-industry transition rates and find that signers have 2 percentage points lower within-industry mobility.[15] Despite the higher wage levels, we find no significant relationship between signing an NC and wage growth or formal training in the cross-section.

There are also notable differences in the types of workers who report having NCs. Signers tend to have characteristics associated with higher earnings: 52 percent hold a bachelor's degree or higher, compared to 42 percent among non-signers, and they score higher on the ASVAB, a standardized test assessing science, math, and language skills that we use as a proxy for cognitive ability. They are also less likely to be unionized and more likely to have negotiated over wages. Compared to non-signers, NC signers are more likely to be male and less likely to identify as Black or Hispanic. We find no significant differences in job satisfaction or firm size across the two groups.

The raw wage premium associated with NCs is partly explained by differences in observables

[14]NLSY97 respondents appear confident in reporting their NC status, which contrasts with contemporaneous findings from Cowgill, Freiberg, and Starr (2024). In a field experiment with job applicants for full-time positions at a large U.S. financial services company, they document that many workers fail to notice or recall NCs in their contracts, particularly when the NC is not made salient at the time of signing.

[15]It is worth noting that within-industry mobility accounts for only about one-third of all job changes for both groups. This finding is consistent with Parent (2000), who shows using NLSY79 that roughly two-thirds of job moves occur across one-digit industries. The frequency of between-industry job mobility is also documented in Neal (1999) and Kambourov and Manovskii (2008). The prevalence of inter-industry transitions, especially among younger and more mobile workers, aligns with the modest job mobility effects we document. Consistent with this reasoning, an NC in our theoretical model allows for job mobility across industries.

(Table E3). Once we control for observable factors such as education, tenure, experience, ASVAB test scores, and the industry and occupation of employment, the estimated cross-sectional wage premium between NC signers and non-signers falls from 21.1 log points to 8.1 log points. In contrast, the estimated relationships between NCs and other labor market outcomes, such as changing one's main employer, wage growth, and formal training remain largely stable as we introduce additional covariates. We arrive at similar results when using the 2019 cross-section, which only consists of new job holders (Table E4). Overall, the conclusions from our descriptive relationships align with those reported in Rothstein and Starr (2022), who study the cross-sectional wage premium earned by NC signers in 2017 using NLSY97 data.

3.2 NCs and Task Content

Existing research suggests non-routine cognitive skills are more easily transferable across jobs, and therefore if firms use NCs to provide transferable skills as our model suggests, we should observe that NCs are used more intensively in jobs requiring non-routine cognitive tasks (Acemoglu and Autor 2011; Gathmann and Schönberg 2010). We test this prediction by comparing the task content of NC signers to non-signers, using pooled cross-sectional data from 2017 to 2021 when these data are available.

We find that NC signers are more likely to regularly read long documents, use problem solving and mathematics in their jobs, and are more likely to manage or supervise other workers (Figure II). Conversely, jobs using NCs are less intensive in short and repetitive or physical tasks. Jobs with NCs also involve less face-to-face contact with people outside of their team – a task most common in service industries such as entertainment, accommodations, and food services. We observe similar relationships at the broader industry or occupational level. Occupations that are more intensive in problem solving, mathematics, reading long documents, and managing others tend to use NCs more often, whereas those involving short and repetitive tasks, physical tasks, or external contact use NCs less often (Figures E1a-E1g). Using O*NET measures at the occupational level, we also find that NCs are used more frequently in jobs requiring non-routine cognitive skills and less frequently in manual skills. The industry-level patterns are similar but tend to exhibit a weaker relationship between task content and NC usage (Figures E1h - E1n). Overall, we interpret these findings as suggestive evidence that NCs are used more often when the skills accumulated on-the-job are more transferable to competing firms.

3.3 NCs and Return to Experience

To further probe whether NCs mitigate the hold-up problem, we examine whether NCs are used more frequently in industries and occupations with high returns on experience. As in Topel (1991),

we think of experience as a proxy for general skills and tenure for firm-specific skills.

$$log(w_{ijt}) = \beta_1 + \beta_2 X_{ijt} + \beta_3 T_{ijt} + \theta_i + \mu_{ij} + \varepsilon_{ijt} \tag{7}$$

We estimate the return on experience for each occupation and industry separately using Equation 7, where $w_{ijt}, X_{ijt}, T_{ijt}$ denote the real hourly wage, experience and tenure of individual i in firm j in year $t \in \{2017, 2019, 2021\}$. θ_i denotes the individual fixed effect, μ_{ij} denotes the worker-firm pair fixed effect, ε_{ijt} represents the random shock component, and β_2 reflects the return to experience. To estimate the returns to tenure and experience, we apply the two-step estimation developed by Topel (1991) and reviewed by Altonji and Williams (2005). The exact details on the estimation procedure are described in Appendix D. We find that there is a significant and positive correlation between the return to experience and mean NC usage by occupation, where we see a 0.8 log point increase in return to experience for every 10% increase in NC usage (Figure E2a). There is also a positive albeit weaker correlation on the industry level, as shown in Figure E2b.[16] These results complement our analysis of task-content to support the theory that NCs are more frequently used when the returns on transferable skills are larger. They also align with the findings in Lavetti, Simon, and White (2020), who use a modified version of Topel's two-step procedure to show that NCs reduce the return on experience and increase the return on tenure among physicians.

3.4 Stacked Difference-in-Differences Approach

To estimate the causal effects of signing an NC, we adopt a stacked difference-in-differences (DiD) research design. This approach is well-suited to settings with staggered treatment timing, where each treated cohort requires its own control group. Rather than relying on a single pool of never-treated units—as in the estimator proposed by Callaway and Sant'Anna (2021)— the stacked DiD approach constructs cohort-specific panel datasets, each with its own treatment and control group, and then combines (or stacks) these datasets to estimate a single regression model with either dynamic or aggregated treatment effects. This design has been applied in several empirical contexts, including firm responses to liability risk (Gormley and Matsa 2011), minimum wage changes (Cengiz et al. 2019), and state-level changes in NC enforceability (Johnson, Lavetti, and Lipsitz 2025; Jeffers 2023).

In our setting, the timing of NC adoption varies across individuals, and treatment is closely tied to job transitions. For each survey year c, we construct a cohort-specific panel dataset that includes

[16] In these plots, we depict the return to experience against mean NC usage by occupations and industries. Each circle represents an occupation/industry and is proportional to its size. The fitted line is also weighted by the size of each occupation/industry.

only those who moved to a new job with known NC status in year c. The treated group consists of workers who report signing an NC for the first time in year c; the control group includes job movers in the same year who never sign an NC during the sample period. Denote G_c as the set of treated individuals in cohort c. We focus on cohorts $c \in \{2015, 2017, 2019, 2021\}$ and restrict to data from 2013–2021 to allow a sufficient post-treatment horizon without assuming long-run within-job stability in NC status.[17] We then stack the data across cohorts and estimate the following event-study specification:

$$Y_{itc} = \alpha_{ic} + \lambda_{tc} + \sum_{k \in \{-6+,-4,0,2,4,6\}} \beta^k d_{i,t+k,c} + \varepsilon_{itc} \tag{8}$$

where Y_{itc} is the outcome of interest for individual i at time t in cohort c, and $d_{i,t+k,c}$ is an event-time indicator equal to one for treated individuals $i \in G_c$ in year $t = c + k$.

For ease of interpretation and to increase precision, we also estimate an aggregated model by collapsing the dynamic indicators into a single post-treatment dummy:

$$Y_{itc} = \alpha_{ic} + \lambda_{tc} + \beta^{Agg} d^{Agg}_{i,t,c} + \varepsilon_{itc} \tag{9}$$

where $d^{Agg}_{i,t,c}$ equals one for treated individuals in all periods $t \geq c$, and β^{Agg} captures the average post-treatment effect. Standard errors are clustered at the individual level.

This design has several advantages. First, the inclusion of individual-by-cohort fixed effects α_{ic} and time-by-cohort fixed effects λ_{tc} ensures that workers treated in different periods are not inappropriately used as controls, addressing the "bad comparison" problem highlighted in the recent DiD literature (e.g., Sun and Abraham 2021; Callaway and Sant'Anna 2021). Second, by design, our approach compares job movers to job movers, addressing a challenge in this context. Because we only observe NC status when workers transition to a new job, any change in treatment status is necessarily tied to job mobility.[18] Since job changes are often associated with changes in wages, training, and other labor market outcomes, using all never-treated workers as a pooled control group would confound the effects of NCs with the effects of job mobility. By defining cohort-specific control groups of other job movers, our design ensures a more appropriate comparison. Finally, the inclusion of individual-cohort fixed effects allows us to control for time-invariant un-

[17] We note that NC status is unobserved for jobs ending prior to 2017. This implies that treated workers in cohort 2015 are necessarily job stayers between 2015 and 2017. We impose the same restriction on the control group. Unobserved pre-2017 NC status also implies that some of the jobs held prior to year c for both the treatment and control group may have had an NC. We confirm our results are not sensitive to restricting to later cohorts that are unaffected by this issue.

[18] This feature is not merely an artifact of survey design—the NLSY97 does not ask about NC status across waves for the same job—but reflects the fact that NCs are typically signed upon starting a new job (Starr, Prescott, and Bishara 2021).

observable characteristics that may influence both selection into NCs and subsequent labor market trajectories.

3.5 Identifying Assumptions

Our identification strategy relies on a parallel trends assumption that rules out selection into treatment based on time-varying unobservable characteristics. In particular, for the event-study coefficients $\hat{\beta}^k$, $k \geq 0$, from equation (8) to be interpreted as the causal effect of signing an NC, we assume that, within each cohort c, NC signers and their matched control group (job movers who never sign an NC during the sample period) would have experienced similar trends in potential outcomes in the absence of signing an NC.

Given that NC status is not randomly assigned but instead reflects joint decisions by workers and firms, this assumption may fail if selection into NCs is based on unobserved characteristics that vary over time. While the inclusion of individual fixed effects α_{ic} controls for selection based on time-invariant unobservables (e.g., inherent ability), time-varying unobserved shocks could still bias our estimates. We now consider two types of time-varying unobservable factors that could threaten identification and explain why these are unlikely to bias our results.

The first concern is selection based on unobserved job characteristics. Certain types of jobs may be more likely to include NCs; for instance, jobs involving access to client relationships or managerial responsibilities. If these jobs also have inherently higher wages, our wage estimates may be upward biased. To address this concern, we re-estimate our models controlling for industry, occupation, and firm size decile fixed effects – characteristics strongly linked to firm productivity (Manning 2013). If unobserved job characteristics were driving the results, we would expect these controls to attenuate the estimated wage effects. Reassuringly, our results are nearly unchanged. We also show that NC signers are not disproportionately moving into jobs involving greater task complexity. For example, they are not significantly more likely to begin work involving management, problem-solving, or external-facing client work. These robustness checks suggest that our results are not driven by NC signers sorting into higher productivity firms.

The second concern is selection based on unobserved worker characteristics that change over time. For example, workers may experience positive shocks – such as earning a new credential, receiving strong performance reviews, or gaining reputation – that are observed by employers but not the econometrician. Such events could potentially increase the likelihood of sorting into NC-covered jobs (e.g., through promotions or matching with more selective employers), while also increasing wages independently of NCs. In this case, the NC treatment indicator d_{itc} would be correlated with the error ε_{itc}, biasing our wage estimates upward.

While we cannot directly observe these time-varying shocks, we present several pieces of evi-

dence suggesting that they are unlikely to bias our results. First, if such shocks occurred, we would expect NC signers to experience systematically different trends in outcomes prior to signing, such as faster wage growth. However, our event study plots and group-level means show no such pre-trends (Figure III). Mean tenure is relatively flat for both groups, and while wage growth is slightly higher for NC signers four years prior to signing, it is slightly lower two years prior to signing. In our event studies, these differences are not significant, and in each case we formally test the joint null that all pre-treatment coefficients are zero, following Borusyak, Jaravel, and Spiess (2024), and typically find large p-values associated with the F-test. Second, if selection reflected improved job match quality, we would expect observable shifts in job composition. The fact that NC signers are not sorting into jobs requiring more complex tasks and that our results hold when controlling for job-level characteristics such as occupation, industry, and firm size, suggests that this form of selection is not driving our estimates. As further robustness, we also try alternative empirical specifications, such as the standard two-way fixed effects estimator or using the later-treated as the control group. We leave the details of these exercises and their results to Section 4.5.3, but note here that our results are highly robust in each case.

4 Results

4.1 The Effect of NCs on Wages and Job Mobility

Figure IV presents our main event-study estimates, with results reported separately for log wages and job tenure. NCs significantly raise wages upon signing, with the effect diminishing over time: wages increase by 9.9 log points in the year of signing and remain 4.7 log points higher six years later. This pattern mirrors descriptive findings in Shi (2023) among executives and suggests that while NC signers receive an immediate wage premium, their subsequent wage growth is slightly slower, averaging about 1 log point less per year. Signing an NC also leads to longer job tenures: we estimate an initial increase of 0.27 years (roughly 3 months), growing to 0.49 years (6 months) six years after signing.[19] These results highlight the retention effect of NCs, in line with our theoretical model.

Next, Table II reports aggregated estimates from equation (9), which improves statistical precision and facilitates the comparison of results across multiple outcome variables. NCs are associated with a 9.4 log point increase in wages, a 0.29 year increase in job tenure, and a 3.6 percentage point

[19]The positive effect on job tenure at time 0 reflects that NC signers had shorter pre-signing tenures but similar tenures to non-signers at time 0 by construction (Figure III). Our preferred estimate is the pooled post-period coefficient of 0.287 which is similar in magnitude to the event-study coefficient at time 0. We achieve a very similar aggregated point-estimate when using time 0 as the reference period, as opposed to time -2. Overall, despite similar average tenures across groups in the post-treatment period, we find that NCs reduce mobility because NC signers were more mobile than non-signers in the pre-treatment period.

decline in the biannual probability of changing main employers. The wage effect is statistically significant at the 1% level and the mobility effects are significant at the 5% level. We find no significant effect on within-industry job mobility, possibly due to small sample sizes: only 150 NC signers change employers after signing, and fewer than 50 do so within the same industry, making inference difficult. The magnitude of these effects is economically meaningful. A 9.4 log point wage gain is roughly equivalent to the return to an additional year of education (Card 1999), or two years of typical wage growth in our sample. In line with the magnitudes in Lavetti, Simon, and White (2020) and Marx, Strumsky, and Fleming (2009), the decline in job mobility corresponds to a 12% reduction relative to the 2017 mean and the increase in tenure represents a 6% rise.

4.2 The Effect of NCs on Formal Training and Other Outcomes

Beyond wages and mobility, we also investigate the effect of signing an NC on other labor market outcomes, including hours, job satisfaction, and formal training (Table III).[20] The most noteworthy result is the absence of any significant effect of NCs on the incidence or intensity of formal training programs. We test the incidence of any training and employer-provided training, finding a null effect. Conditional on receiving formal training, we also test the duration of that training in weeks, hours per week and total hours spent in training, and an indicator for completing the training program. We find a positive coefficient for duration and completion rate of training but a negative effect on hours spent in training, and none of the coefficients are significant at conventional levels. These results are in contrast to Starr (2019), who finds an increase from no enforcement of NCs to mean enforcement is associated with a 14% increase in training, which tends to be firm-sponsored.

That NCs do not appear to increase training for workers appears at odds with our theoretical prediction that NCs should raise incentives for firm-provided investment in human capital. However, several lines of reasoning caution against this conclusion. First, the NLSY97 questionnaire for these training variables asks about a "school, course or training program", suggesting these variables capture more formal types of firm-provided training. Since only 10% of workers report receiving training of this type in 2017, this measure likely overlooks informal or on-the-job learning (e.g., Mincer 1962).

Second, we find that there is a weak relationship between formal training, wages, and wage growth (Table E5). Workers who report formal training earn slightly higher wages, but most of this wage premium is explained by worker observables (columns 1-3). Importantly, if formal training programs were an important source of human capital growth we would expect a positive relationship between their incidence and wage growth. However, we find an insignificant relationship across all of our specifications (columns 4-6). The minimal relationship between the incidence

[20] To conserve space, we report treatment effects aggregated across all post-treatment periods. Results showing flat pre-trends for the outcomes in Table III are available upon request.

of formal training programs and wage growth aligns with findings by Veum (1995), who studies these relationships using data from older cohorts in the NLSY79. Taken together, the null effect of NCs on formal training does not rule out the possibility that NCs facilitate investment in worker productivity through other channels.

Third, while formal training rates do not significantly differ between groups, the fact that NC signers continue to earn higher wages even six years after signing suggests that NCs lead to productivity-enhancing investment. If NCs do not encourage investment but instead function purely through a compensating differential, then we would expect to see a one-time shift in the timing of compensation: workers would receive higher earnings in the period of signing, but lower earnings in subsequent periods (Corollary 2). The persistence of a 4.7 log point wage premium six years after signing suggests that NCs raise worker productivity.[21]

4.3 Movers vs. Stayers

In our theoretical framework, the wage effects of NCs depend on whether workers are expected to stay or leave. Patient workers who are expected to remain with the firm receive greater investment and, as a result, higher contracted wages. By contrast, when firms anticipate a higher probability of quitting (low λ), they scale back investment and wages within the firm are lower. Moreover, if such workers do leave, they can only access the non-industry outside option (v). The theory thus predicts that wage gains from NCs should be concentrated among stayers rather than movers.

To test this prediction, we split treated workers into two groups by their post-treatment job mobility. A treated-stayer is defined as a worker who signs an NC at time c and remains in that same job in all subsequent observed periods $t > c$. A treated-mover signs an NC at time c but changes employers at least once in the post-period. Both groups are compared to a common control group consisting of never-treated workers who also changed jobs at time c. Our results are robust to instead using mobility-matched controls (e.g., comparing treated-movers to control-movers and treated-stayers to control-stayers).[22]

Figure V shows that the positive wage effects of NCs are driven primarily by stayers. Treated-stayers receive a 12.1 log point wage premium at signing and still earn 3.3 log points more six years later. In contrast, treated-movers earn 4.3 log points more than the control group at the time of signing and 4.4 log points more six years later, but these estimates are imprecise and there is no clear break at the point of treatment. These findings reinforce the interpretation of our main

[21] Our model attributes increased worker productivity to employer-provided investments. We cannot fully rule out that NCs raise productivity through worker-initiated investments, such as greater effort. That said, we do not detect significant effects on hours worked, a commonly used proxy for worker effort.

[22] We acknowledge that in this analysis, we are conditioning on a post-treatment outcome variable. In the remaining heterogeneity analyses, we split based on time-invariant characteristics or pre-sample outcome variables.

results, showing that the persistent NC wage premium is not driven by treated workers who exit to better-paying opportunities but rather by those who remain.

4.4 Heterogeneity by Worker Demographics

While NCs raise wages on average, an important question is whether these positive effects extend to all types of workers. This inquiry is particularly pertinent given recent policy discussions and enacted bans on NCs for low-wage workers (e.g., Lipsitz and Starr 2022; Young 2021). Our model suggests that NCs are most beneficial in situations where firm-sponsored, industry-specific investments are substantial. However, our model also predicts a more exploitative scenario: workers with higher discount factors and lower returns on firm-provided investment may accept NCs even in the absence of significant firm investment. This dynamic is more likely to apply to lower-wage workers, who may be more financially constrained and therefore value current income more. Motivated by these theoretical insights, we next explore how the effects of signing an NC vary across different worker subgroups.

We categorize workers by educational attainment, sex, race, average hours worked per week, and pre-sample wages (above or below median wages in 2011). The estimated wage increases range from 7.3 log points for Black/Hispanic workers to 12.4 log points for workers with above-median 2011 wages, and all effects are significantly positive (Figure VI). Similarly, the estimated tenure effects, while consistently positive, reveal notable differences across subgroups. For instance, college-educated workers experience a 0.51-year increase in tenure compared to 0.20 years for those without a college degree, and high-wage workers see a 0.55-year increase versus 0.23 years for low-wage workers. On the surface, this analysis offers little support for the idea that NCs systematically disadvantage certain subgroups.

However, these aggregate wage effects mask patterns in wage trajectories over time. We delve deeper by conducting our dynamic event study for each worker subgroup, revealing that while NCs lead to higher wage levels for all worker types over the first six years, notable differences emerge in their relative wage trajectories (Figure VII). Workers with less than a college education, those who identify as Black or Hispanic, those with low pre-sample wages, and those who typically work less than 40 hours per week all experience an upfront wage premium close to the average. Yet, this initial gain is followed by negative wage growth relative to their No-NC counterparts. In fact, for Black/Hispanic, low-income workers, and those who typically work less than 40 hours per week, the wage point estimate turns negative by year six, though null effects cannot be rejected at conventional levels. Conversely, college-educated, non-Black/Hispanic, and high-wage workers exhibit similar upfront wage premia but maintain flat or even increasing wage-experience profiles, as among the physicians studied in Lavetti, Simon, and White (2020).

For these higher-skill workers, systematically higher wages combined with stable or even increasing subsequent wage growth strongly contradicts the notion that NCs exploit them or trap them in lower-paying jobs. In contrast, the wage dynamics for workers with lower wages, less education, and those belonging to racial minority groups paint a different picture. The pattern of an upfront wage premium followed by negative wage growth is more consistent with limited investments and higher initial wages reflecting a compensating wage differential. As such, our findings strongly suggest that policy interventions such as bans on NCs should be carefully tailored, perhaps specifically targeting lower-wage workers.

4.5 Robustness

4.5.1 Addressing Potential Omitted Variable Bias

Our primary identifying assumption is that the error terms in our core regressions are uncorrelated with NC signage, conditional on our control variables. This identifying assumption rules out unobserved, time-varying shocks that are simultaneously correlated with NC signage and labor market outcomes, such as wages and job mobility. We see two potential violations of this identifying assumption. First, workers who sign NCs may join higher-productivity firms that would have paid higher wages even without an NC. Second, workers may experience shocks to their real or perceived productivity that lead them to sign an NC and that would have resulted in higher wages even without one. These time-varying unobservables introduce a potential omitted variables problem that could generate upward bias in our estimated wage effects. While our data limits our ability to control for these unobservables directly, we conduct several robustness checks to address these potential biases.

Our first robustness check is to re-estimate our aggregate effects from equation (9) for wages, tenure and the job mobility indicator, progressively adding fixed effects for industry, occupation, and firm size decile. Industry and occupation fixed effects account for the possibility that NCs are more prevalent in certain high-paying sectors. Firm size is a well-established proxy for firm productivity in labor market models (e.g., Manning 2013), and in results available upon request, we confirm its strong correlation with wages in the NLSY97 data. We find that controlling for these job-level characteristics has minimal impact on our results (Figure E3). The estimated wage effect decreases by about 1 log point when controlling for both industry and occupation fixed effects. Similarly, our job mobility estimates remain robust to these additional industry and occupation controls; the tenure effect shifts modestly from 0.29 to 0.26 years, and the impact on the rate at which workers change main employers changes from -3.6 to -3.2 percentage points. We arrive at similar conclusions when we also add firm size as a control variable, though standard errors rise because firm size data is missing for approximately one-third of observations in our sample.

As a second robustness exercise, we compare how the task content of a worker's job changes for NC signers relative to non-signers. Starting in 2017, the NLSY97 asks workers about the task content of their jobs, from which we construct indicator variables based on whether the job involves frequent short and repetitive tasks, physical tasks, management or supervision, problem solving, math, reading long documents, or in-person contact with people other than co-workers or supervisors. These data prevent us from examining whether signing an NC results in systematic upgrading of skills in the long-run, but we can ask whether NC signers are immediately assigned more sophisticated tasks – a potential omitted variable. Using just the 2019 and 2021 cohorts to capture short-term effects, we estimate our aggregate stacked difference-in-differences regression using each of the task-indicator variables as outcome variables. The estimated coefficients capture the extent to which workers moving into jobs with an NC change tasks in a way that is systematically different from those moving into a job without an NC. Our estimates show no significant differences between NC signers and non-signers in terms of changes in task content (Figure E4).[23]

If the wage premium we observe is driven by NC signers systemically moving into higher-productivity jobs, we would expect to see job level controls, such as the sector and size of the firm, mitigating our estimates. Similarly, if NC signage is correlated with between-firm promotions into higher paying jobs, such as those induced by worker productivity shocks, we would expect to see systematic changes in the types of tasks the worker is performing. The fact that neither of these predictions bear out in the data, coupled with the observation that NC signers have similar pre-signage wage growth, give us confidence that our baseline estimates are not biased by unobserved, time-varying characteristics positively associated with both wages and NC status.

4.5.2 Data Limitations on NC Usage

The structure of the NC questions in the NLSY97 survey requires us to place assumptions on the timing of NC signage. First, since we only observe NC status starting in 2017 and do not observe changes to NC status within the same job, we assume that the NC was signed at the beginning of the employment relationship. This assumption imposes that NC status does not change with job tenure. A violation of this assumption would arise if NCs were frequently signed later on in an employment relationship – for instance, if NC signage is tied to promotions within the same job. Existing research by Starr, Prescott, and Bishara (2021) indicates that such instances are rare, with the vast majority of NCs signed at the beginning of the employment relationship.

A related data limitation is the absence of NC status in the pre-period for cohorts 2015 and 2017. To formally assess the impact of this limitation, we conduct a robustness check by sequentially removing these cohorts from our sample. Specifically, we re-calculate our aggregate estimates separately using (1) only years 2015-2021 and cohorts 2017-2021 and (2) only years

[23] We arrive at similar conclusions when using the standard TWFE model, given by equation 11.

2017-2021 and cohorts 2019-2021. The latter restriction is particularly valuable as it ensures we only utilize observations where NC status is directly observed both before and after a job move, providing more transparent estimates.

The results of this exercise are reassuring. We find that our estimated wage effects remain largely unchanged; if anything, they tend to increase slightly in magnitude as we remove the 2015 and 2017 cohorts (Figure E5). Similarly, the point estimates for job tenure and job mobility are similar across these restricted samples, though standard errors rise as expected when dropping cohorts.[24] This consistency demonstrates that our findings are not driven by measurement error in NC status and also indicates the stability of our results across different cohorts. Finally, it is important to acknowledge that NCs are often bundled with other post-employment restrictions, such as non-disclosure agreements (Balasubramanian, Starr, and Yamaguchi 2024). To the extent that this bundling occurs within our data, our estimates should be interpreted as identifying the causal effects of this broader set of restrictions, rather than NCs in isolation.

4.5.3 Alternative Estimation Methods: Later-Treated and TWFE

In this section we explore the sensitivity of our results to alternative estimation methods. First, we explore the implications of redefining our control group as the "later-treated" rather than the "never-treated" workers. Specifically, we re-estimate our stacked event-study specification given by equation (8), but now define the control group for each cohort c as job-movers in the same year who do not yet have an NC but go on to sign one in a later survey wave ($t > c$). This specification compares NC signers to future NC signers, leveraging only the timing of contract adoption. The identifying assumptions remain similar to those in our baseline specification.

Relative to our baseline stacked event-study specification, we observe a larger initial wage effect when using later-treated controls, followed by more negative wage growth (Figure E6). The negative post-signing wage trajectory is expected: individuals in the control group are themselves on track to sign NCs and thus likely to experience wage increases around their own signing dates. As a result, the treated group's wage growth appears weaker after time zero relative to our baseline specification. While this feature of the later-treated control group complicates interpretation of post-treatment effects, the strong initial wage gain reinforces that our baseline estimates are not driven by unobserved differences between NC signers and non-signers.

We also estimate two specifications using a standard two-way fixed effects (TWFE) approach, which departs from our stacked design and includes all workers, not just job movers. Specifically,

[24] We also arrive at similar conclusions when running our stacked wage event study using survey years strictly past 2010 and cohorts strictly past 2015, yielding a four year effect window. The initial wage premium is 0.094, which falls to 0.083 in the four years past NC signage, with all coefficients significant at the 5% level.

we estimate:

$$w_{it} = \alpha_i^a + \lambda_t^a + \beta^{TWFE-a} d_{it}^a + \varepsilon_{it}^a, \tag{10}$$

$$w_{it} = \alpha_i^{na} + \lambda_t^{na} + \beta^{TWFE-na} d_{it}^{na} + \varepsilon_{it}^{na} \tag{11}$$

where w_{it} denotes log wages, α_i^j are individual fixed effects, λ_t^j are year fixed effects, and ε_{it}^j is the error term for specification $j \in \{a, na\}$. The variable d_{it}^a is an absorbing treatment indicator equal to one from the first observed year of NC signage onward. In contrast, d_{it}^{na} is a non-absorbing treatment indicator, capturing only current NC status and allowing for transitions into and out of NC status within individuals. We limit the *na* model to the 2017–2021 period, when both entry and exit from NC jobs can be directly observed.

The point estimates from the standard TWFE model with the absorbing treatment indicator yield results that closely mirror our baseline results. The estimated aggregate wage effect is 9.5 log points, virtually identical to our baseline estimate of 9.4 log points (Figure E7a). Estimating the parameters from the TWFE model with the non-absorbing treatment indicator yields a smaller but still statistically significant average wage effect of 5.8 log points (Figure E7b). In both sets of analyses, the point estimates for higher wage workers are larger than for lower wage workers, but confidence intervals overlap, preventing us from making strong claims about heterogeneity.

5 Conclusion

Economists have long been interested in the factors that promote human capital development, but the market for employer-provided training suffers a well-known failure: employers are reluctant to provide transferable skills if they must later pay for the worker's increased productivity (e.g., Becker 1962; Acemoglu and Pischke 1999). Our primary theoretical contribution is an asymmetric information model which shows that NCs can mitigate this problem by encouraging investment in industry-specific human capital investments, but at the cost of generating ex-post allocative inefficiencies. This model predicts that NCs increase wage growth due to human capital appreciation, which matches the data well for the typical NC signer who is college-educated. We also present a liquidity constraints model to explain the usage of NCs among lower-wage workers. For liquidity constrained workers with lower returns on investment, NCs will be used to raise firm profits, even if the resulting increase in investment is minimal. The theory predicts these workers receive a compensating wage differential for accepting restricted mobility but will not experience wage growth due to human capital appreciation. Therefore, while NCs may be mutually beneficial to workers and firms across various segments of the labor market – providing a theory for their broad usage – the effects of NCs on the distribution of economic surplus can vary substantially.

Relative to previous research that has focused on the causal effects of NC regulation, which generally finds that higher enforcement leads to lower wages (e.g., Johnson, Lavetti, and Lipsitz 2025), our study is one of the first to use a quasi-experimental design to isolate the causal effects of signing an NC itself. As expected, we find that signing an NC lowers job mobility, raising job tenures by 6% and lowering rates of job-to-job transitions by 12%. These point estimates align with those in Lavetti, Simon, and White (2020), who use similar empirical methods to study the effects of signing NCs for physicians. In contrast to the sizable negative wage effects of higher enforcement, we find sizable positive wage effects of signing an NC. Using data on NC usage, future work could investigate the mechanisms driving the negative wage effects of higher enforcement. A starting point would be to see if there is a relationship between NC enforcement and usage, analogous to prior work studying the relationship between Right to Work laws and unionization.

Author Affiliations

- Bhargav Gopal: Assistant Professor at Smith School of Business, Queen's University.
- Xiangru Li: Doctoral student at Columbia University, Department of Economics.
- Luke Rawling: Doctoral student at Queen's University, Department of Economics.

References

Acemoglu, Daron and David Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings." *Handbook of Labor Economics*. Vol. 4. Elsevier, 1043–1171.

Acemoglu, Daron and Jörn-Steffen Pischke (1999). "The structure of wages and investment in general training." *Journal of political economy* 107.3, 539–572.

Altonji, Joseph G. and Nicolas Williams (2005). "Do Wages Rise with Job Seniority? A Reassessment." *ILR Review* 58.3, 370–397.

Balasubramanian, Natarajan, Evan Starr, and Shotaro Yamaguchi (2024). "Employment restrictions on resource transferability and value appropriation from employees." *Strategic Management Journal* 45.12, 2519–2547.

Becker, Gary S. (1962). "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70.5, 9–49.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024). "Revisiting Event-Study Designs: Robust and Efficient Estimation." *Review of Economic Studies* 91.6, 3253–3285.

Callaway, Brantly and Pedro H.C. Sant'Anna (2021). "Difference-in-Differences with multiple time periods." *Journal of Econometrics* 225.2, 200–230.

Card, David (1999). "The causal effect of education on earnings." *Handbook of labor economics* 3, 1801–1863.

Cengiz, Doruk et al. (2019). "The Effect of Minimum Wages on Low-Wage Jobs*." *The Quarterly Journal of Economics* 134.3, 1405–1454.

Che, Yeon-Koo and Donald B Hausch (1999). "Cooperative investments and the value of contracting." *American Economic Review* 89.1, 125–147.

Cowgill, Bo, Brandon Freiberg, and Evan Starr (2024). "Clause and Effect: Theory and Field Experimental Evidence on Noncompete Clauses."

Deming, David J. (2017). "The Growing Importance of Social Skills in the Labor Market*." *The Quarterly Journal of Economics* 132.4, 1593–1640.

Feess, Eberhard and Gerd Muehlheusser (2003). "Transfer fee regulations in European football." *European Economic Review* 47.4, 645–668.

Freeman, Richard B. (1984). "Longitudinal Analyses of the Effects of Trade Unions." *Journal of Labor Economics* 2.1, 1–26.

Freeman, Richard B. and James Medoff (1985). "*What Do Unions Do?* By Richard B. Freeman and James L. Medoff." *ILR Review* 38.2, 244–263.

Gathmann, Christina and Uta Schönberg (2010). "How general is human capital? A task-based approach." *Journal of Labor Economics* 28.1, 1–49.

Gopal, Bhargav, Xiangru Li, and Luke Rawling (2026). "Non-Compete Agreements and Bargaining." *AEA Papers and Proceedings*.

Gormley, Todd A. and David A. Matsa (2011). "Growing Out of Trouble? Corporate Responses to Liability Risk." *Review of Financial Studies* 24.8, 2781–2821.

Gottfries, Axel and Gregor Jarosch (2023). "Dynamic Monopsony with Large Firms and Noncompetes." w31965. National Bureau of Economic Research, w31965.

Gregg Lewis, H. (1986). "Chapter 20 Union relative wage effects." *Handbook of Labor Economics*. Vol. 2. Elsevier, 1139–1181.

Guimaraes, Bernardo, Joao Paulo Pessoa, and Vladimir Ponczek (2025). "Noncompete Agreements, Wages and Efficiency: Theory and Evidence from Brazilian Football." *The Economic Journal*, ueaf103.

Harris, Milton and Bengt Holmstrom (1982). "A Theory of Wage Dynamics." *The Review of Economic Studies* 49.3, 315.

Hashimoto, Masanori (1981). "Firm-specific human capital as a shared investment." *The American Economic Review* 71.3, 475–482.

Heckman, James J., Lance J. Lochner, and Petra E. Todd (2006). "Chapter 7 Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond." *Handbook of the Economics of Education*. Vol. 1. Elsevier, 307–458.

Jäger, Simon, Suresh Naidu, and Benjamin Schoefer (2025). "Collective bargaining, unions, and the wage structure: an international perspective." *Handbook of Labor Economics*. Vol. 6. Elsevier, 229–372.

Jeffers, Jessica S (2023). "The Impact of Restricting Labor Mobility on Corporate Investment and Entrepreneurship." *The Review of Financial Studies* 37.1. Ed. by Manju Puri, 1–44.

Johnson, Matthew S., Kurt Lavetti, and Michael Lipsitz (2025). "The Labor Market Effects of Legal Restrictions on Worker Mobility." *Journal of Political Economy* 133.9, 2735–2793.

Johnson, Matthew S. and Michael Lipsitz (2022). "Why Are Low-Wage Workers Signing Noncompete Agreements?" *Journal of Human Resources* 57.3, 689–724.

Kambourov, Gueorgui and Iourii Manovskii (2008). "RISING OCCUPATIONAL AND INDUSTRY MOBILITY IN THE UNITED STATES: 1968–97*." *International Economic Review* 49.1, 41–79.

Kini, Omesh, Ryan Williams, and Sirui Yin (2021). "CEO Noncompete Agreements, Job Risk, and Compensation." *The Review of Financial Studies* 34.10. Ed. by David Denis, 4701–4744.

Lavetti, Kurt, Carol Simon, and William D. White (2020). "The Impacts of Restricting Mobility of Skilled Service Workers: Evidence from Physicians." *Journal of Human Resources* 55.3, 1025–1067.

Lawrance, Emily C. (1991). "Poverty and the Rate of Time Preference: Evidence from Panel Data." *Journal of Political Economy* 99.1, 54–77.

Lazear, Edward P (1979). "Why is there mandatory retirement?" *Journal of political economy* 87.6, 1261–1284.

Lemieux, Thomas (1998). "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Nonrandom Selection." *Journal of Labor Economics* 16.2, 261–291.

Lipsitz, Michael and Evan Starr (2022). "Low-Wage Workers and the Enforceability of Noncompete Agreements." *Management Science* 68.1, 143–170.

MacLeod, W Bentley and James M Malcomson (1993). "Investments, holdup, and the form of market contracts." *The American Economic Review*, 811–837.

Manning, Alan (2013). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.

Marx, Matt, Deborah Strumsky, and Lee Fleming (2009). "Mobility, Skills, and the Michigan Non-Compete Experiment." *Management Science* 55.6, 875–889.

Meccheri, Nicola (2009). "A note on noncompetes, bargaining and training by firms." *Economics Letters* 102.3, 198–200.

Mincer, Jacob (1962). "On-the-Job Training: Costs, Returns, and Some Implications." *Journal of Political Economy* 70.5, 50–79.

Neal, Derek (1999). "The Complexity of Job Mobility among Young Men." *Journal of Labor Economics* 17.2, 237–261.

Pakes, Ariel and Shmuel Nitzan (1983). "Optimum contracts for research personnel, research employment, and the establishment of rival enterprises." *Journal of labor economics* 1.4, 345–365.

Parent, Daniel (2000). "Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics." *Journal of Labor Economics* 18.2, 306–323.

Posner, Eric A, George G Triantis, and Alexander J Triantis (2004). "Investing in human capital: The efficiency of covenants not to compete." *U Chicago Law & Economics, Olin Working Paper* 137, 01–08.

Potter, Tristan, Bart Hobijn, and André Kurmann (2024). "On the inefficiency of non-competes in low-wage labour markets." *Economica* 91.362, 446–496.

Rothstein, Donna and Evan Starr (2022). "Noncompete agreements, bargaining, and wages." *Monthly Labor Review*.

Rubin, Paul H. and Peter Shedd (1981). "Human Capital and Covenants Not to Compete." *The Journal of Legal Studies* 10.1, 93–110.

Shavell, Steven (1984). "The Design of Contracts and Remedies for Breach." *The Quarterly Journal of Economics* 99.1, 121.

Shi, Liyan (2023). "Optimal regulation of noncompete contracts." *Econometrica* 91.2, 425–463.

Spier, Kathryn E and Michael D Whinston (1995). "On the efficiency of privately stipulated damages for breach of contract: entry barriers, reliance, and renegotiation." *The RAND Journal of Economics*, 180–202.

Starr, Evan (2019). "Consider This: Training, Wages, and the Enforceability of Covenants Not to Compete." *ILR Review* 72.4, 783–817.

Starr, Evan P., J.J. Prescott, and Norman D. Bishara (2021). "Noncompete Agreements in the US Labor Force." *The Journal of Law and Economics* 64.1, 53–84.

Sun, Liyang and Sarah Abraham (2021). "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics* 225.2, 175–199.

Topel, R. H. and M. P. Ward (1992). "Job Mobility and the Careers of Young Men." *The Quarterly Journal of Economics* 107.2, 439–479.

Topel, Robert (1991). "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy* 99.1, 145–176.

Veum, Jonathan R. (1995). "Sources of Training and Their Impact on Wages." *Industrial and Labor Relations Review* 48.4, 812.

Young, Samuel G. (2021). "Noncompete Clauses, Job Mobility, and Job Quality: Evidence from a Low-Earning Noncompete Ban in Austria." *SSRN Electronic Journal*.

Table I: Respondent Characteristics by NC Status in 2017

	NC	No NC	Difference	P Value	N: NC	N: No NC
Job Mobility						
Tenure (Yrs)	5.24	5.11	0.12	0.50	699	4185
1(Main Job Mobility btwn 2017 and 2019)	0.28	0.32	-0.04	0.05	705	4263
1(Within-Industry Job Mobility btwn 2017 and 2019)	0.10	0.12	-0.02	0.08	686	4176
Wages and Wage Growth						
Log(Starting Wage)	2.94	2.76	0.19	0.00	705	4263
Log(Wage in 2017)	3.21	3.00	0.21	0.00	705	4263
$Log(Wage_{2017}) - Log(Wage_{2015})$	0.13	0.12	0.02	0.22	628	3778
$Log(Wage_{2019}) - Log(Wage_{2017})$	0.11	0.10	0.01	0.56	632	3753
Demographics						
Age	35.03	34.96	0.07	0.25	705	4263
1(Male)	0.58	0.50	0.08	0.00	705	4263
1(High School Degree or Higher)	0.89	0.86	0.03	0.01	699	4224
1(Bachelors Degree or Higher)	0.52	0.42	0.10	0.00	699	4224
ASVAB Percentile	57.50	52.06	5.44	0.00	582	3473
1(Black)	0.14	0.16	-0.02	0.13	705	4263
1(Hispanic)	0.11	0.13	-0.01	0.33	705	4263
Wage Bargaining and Negotiation						
1(Possible to Keep Previous Job)	0.46	0.45	0.01	0.74	304	1848
1(Negotiate Job Offer)	0.40	0.33	0.08	0.02	249	1454
Training						
1(Received Some Training)	0.09	0.11	-0.02	0.12	705	4263
1(Received Training Run by Employer)	0.01	0.03	-0.01	0.03	705	4263
1(Received On-Site Training by Non-Employer)	0.01	0.01	0.00	0.64	705	4263
1(Employer Paid for Training)	0.06	0.08	-0.02	0.08	705	4263
1(Employer Paid for Mandatory Training)	0.03	0.04	-0.01	0.26	705	4263
1(Employer Paid for Voluntary Training)	0.03	0.04	-0.01	0.16	705	4263
Other Firm Characteristics						
1(Dislike Job)	0.05	0.06	-0.01	0.57	645	3792
1(Unionized Worker)	0.11	0.16	-0.05	0.00	636	3743
Firm Size	986.28	1134.72	-148.43	0.65	595	3377

Note: The sample includes respondents with valid NC status for the main employer in 2017. All wage variables are measured in terms of real dollars earned per hour. Respondents earning real wages below 3 dollars or above 200 dollars are dropped. The training variables capture whether the respondent received training under any employer in 2017. Means weighted by nationally representative sample weights and p-values from a two-sided t-test are reported. Sample sizes vary due to missing values of the outcome variable. For details on variable definitions, refer to Table E1.

Table II: The Aggregate Effects of Signing an NC on Wages and Job Mobility

	Log Hourly Wages	Tenure	Change Main Emp.	Change Main Emp. Within Industry
	(1)	(2)	(3)	(4)
Treat × Post	0.094***	0.287**	-0.036**	0.006
	(0.022)	(0.126)	(0.016)	(0.018)
Observations	22394	22040	21614	22004
Dependent Variable Mean	2.888	2.692	0.458	0.165
Unique Treated Workers	682	680	681	679
Unique Control Workers	3300	3263	3296	3285
R^2	0.770	0.588	0.598	0.388

Notes: This table reports estimates from stacked difference-in-differences estimation, aggregated over post-treatment years, over a bi-annual sample period of 2013-2021 and using cohorts $c \in \{2015, 2017, 2019, 2021\}$. The treatment group for cohort c are those who we observe first signing an NC in year c. The control group consists of workers who never held an NC during the event window and who also changed jobs between year c and the preceding survey year. Job mobility is defined as changing main employers between the current and preceding survey year. Standard errors are clustered by worker and reported in parenthesis. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

Table III: The Aggregate Effects of Signing an NC on Formal Training and Other Outcomes

	Other Outcomes		Formal Training					
	Log Hours Per Week	Job Dissat.	Any	Employer Paid	Log Weeks	Log Hours Per Week	Log Hours	Completed
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treat × Post	0.034	0.026*	-0.023	-0.022	0.053	-0.309	-0.162	0.103
	(0.024)	(0.016)	(0.017)	(0.015)	(0.197)	(0.232)	(0.301)	(0.064)
Observations	21980	17686	22394	22394	1533	1792	1511	1157
Dependent Variable Mean	3.563	0.072	0.130	0.078	1.901	1.992	3.921	0.904
Unique Treated Workers	681	636	682	682	79	93	78	71
Unique Control Workers	3281	3039	3300	3300	365	413	361	290
R^2	0.511	0.401	0.459	0.455	0.588	0.642	0.609	0.582

Notes: This table reports estimates from stacked difference-in-differences estimation, aggregated over post-treatment years, over a bi-annual sample period of 2013-2021 and using cohorts $c \in \{2015, 2017, 2019, 2021\}$. The treatment group for cohort c are those who we observe first signing an NC in year c. The control group consists of workers who never held an NC during the event window and who also changed jobs between year c and the preceding survey year. Standard errors are clustered by worker and reported in parenthesis. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

Figure I: Timeline of the Model

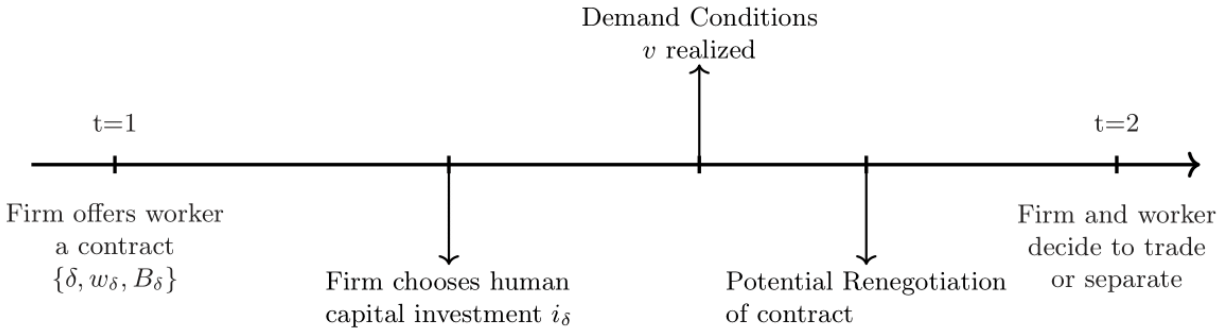
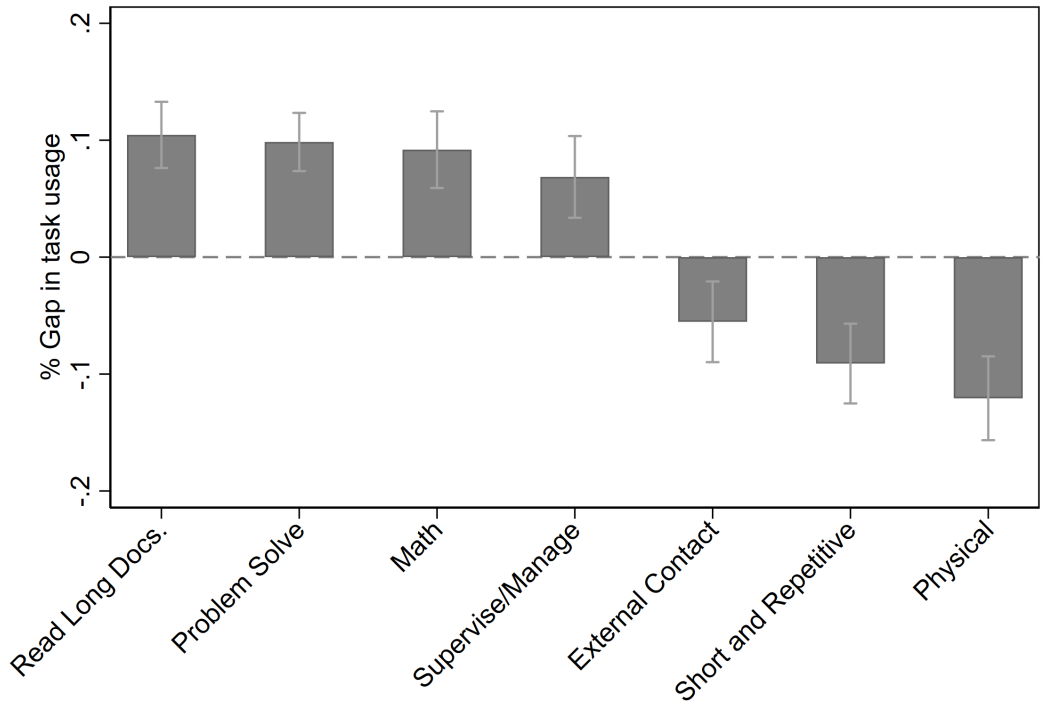


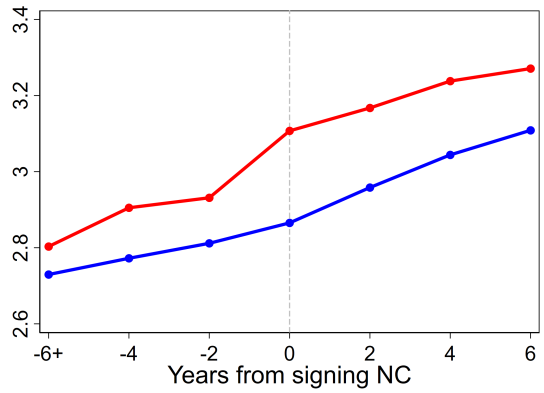
Figure II: Task Content of NC Signers Versus No-NC Signers



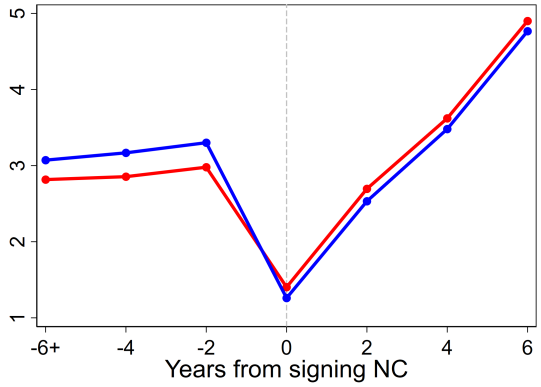
Note: Task content information became available in the NLSY97 data starting in 2017. Sample pools 2017-2021 cross-sections. 95% confidence intervals reported. Standard errors clustered at the individual level.

Figure III: Means Relative to Event Time

(a) Log Wages



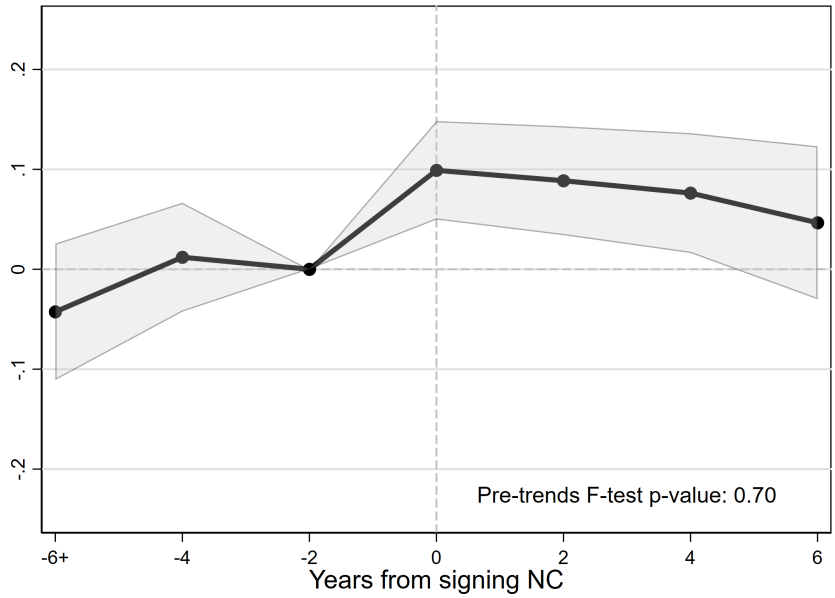
(b) Job Tenure



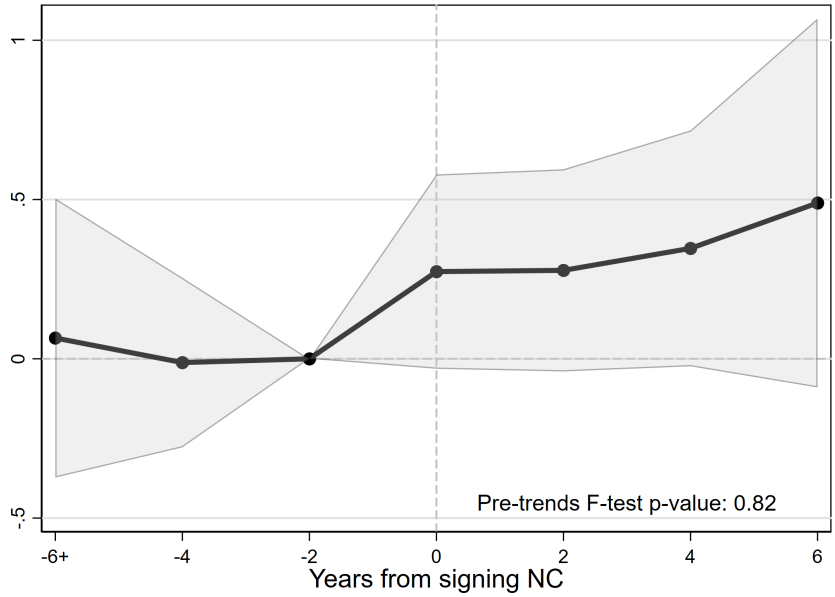
Note: Figure reports means of log wages and job tenure aggregated over all cohorts in the stacked data. See text for data construction. Red lines correspond to NC signers (treatment group) and blue lines to No-NC signers (control group).

Figure IV: The Dynamic Effects of Signing an NC

(a) Log Wages

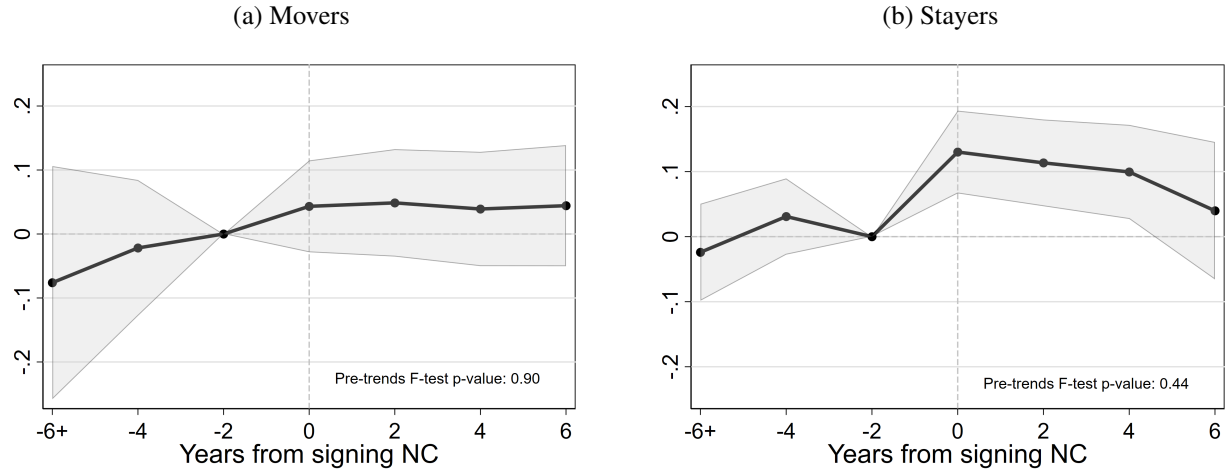


(b) Job Tenure



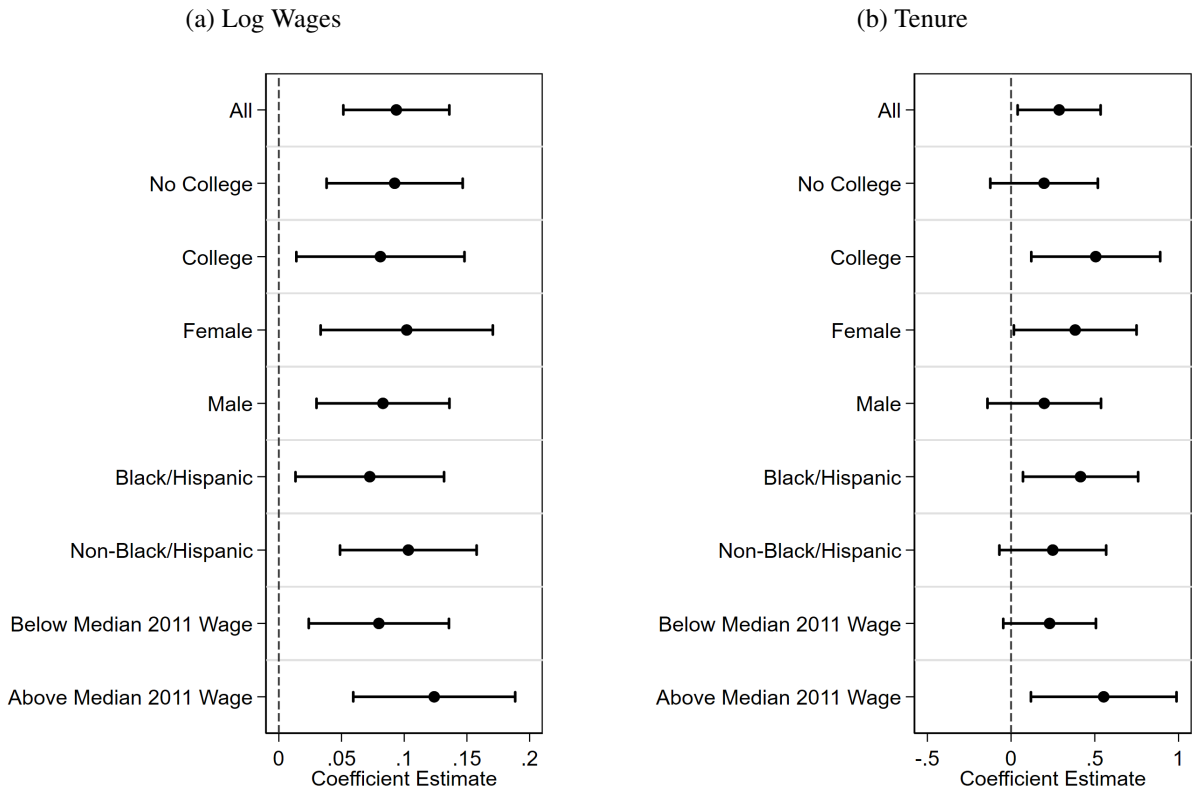
Note: Estimates are from stacked difference-in-differences estimation (equation (8) in the text) over a bi-annual sample period of 2013-2021 and using cohorts $c \in \{2015, 2017, 2019, 2021\}$. The treatment group for cohort c are those who we observe first signing an NC in year c. The control group consists of workers who never held an NC during the event window and who also changed jobs between year c and the preceding survey year. Standard errors are clustered by worker and confidence intervals are reported at the 95% level.

Figure V: The Effect of Signing an NC: Movers vs. Stayers



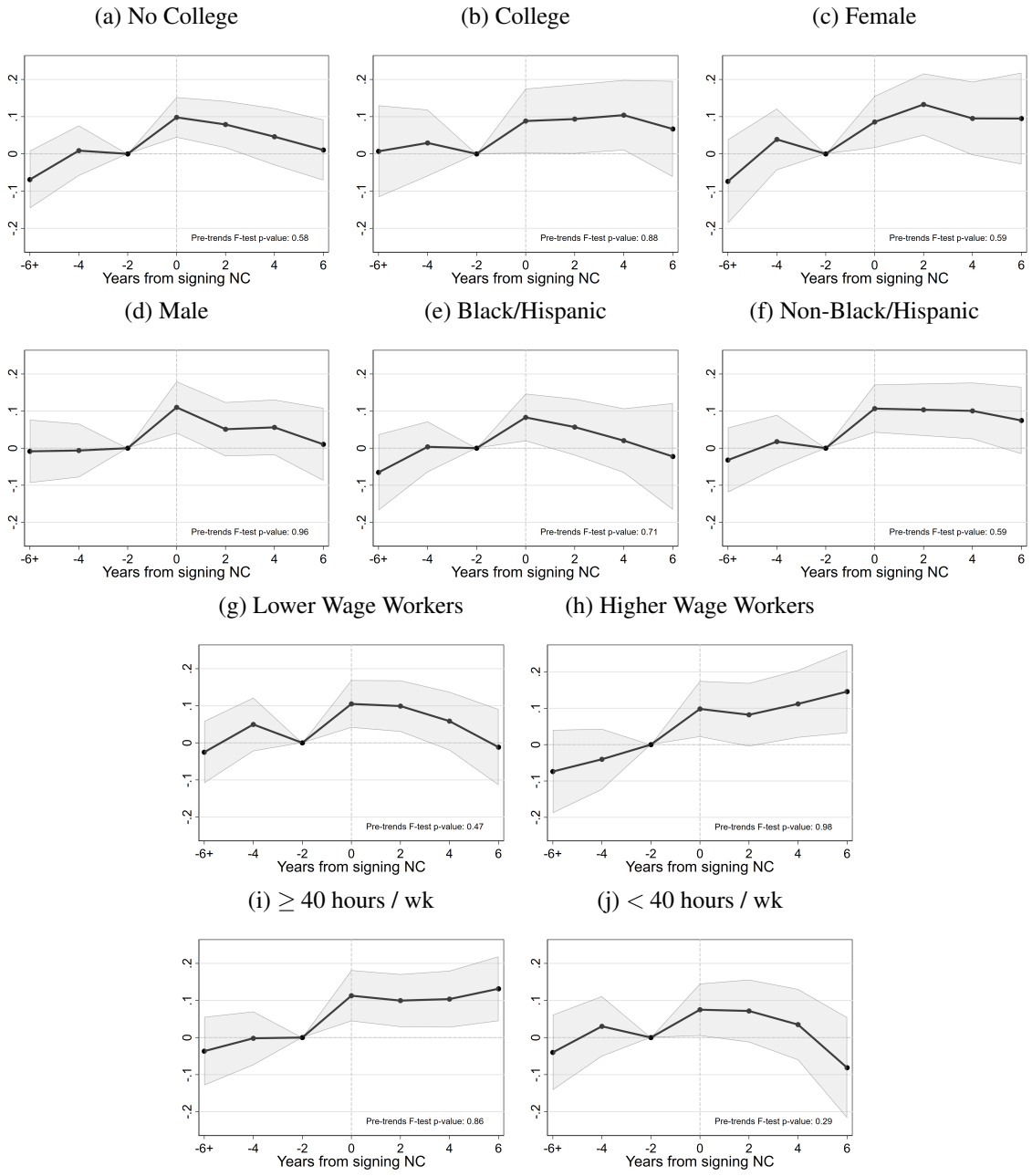
Note: Estimates are from stacked difference-in-differences estimation (equation (8)) using cohorts $c \in \{2015, 2017, 2019, 2021\}$ over a bi-annual sample period of 2013–2021. The treatment group consists of workers who first sign an NC in year c. Movers change main employers at least once post-treatment. Stayers remain in their $t = c$ job throughout the post-period. The control group remains never-treated job movers in year c. Standard errors are clustered by worker. 95% confidence intervals shown.

Figure VI: The Effect of Signing an NC: Heterogeneity Across Worker Groups



Note: Coefficient estimates are from stacked difference-in-differences estimation, aggregated over post-treatment years, using cohorts $c \in \{2015, 2017, 2019, 2021\}$ over a bi-annual sample period of 2013–2021. The treatment group consists of workers who first sign an NC in year c. The control group consists of never-treated workers who changed jobs in year c. Standard errors clustered by worker. 95% confidence intervals shown.

Figure VII: The Dynamic Wage Effects of Signing an NC by Subgroup



Note: Estimates are from stacked difference-in-differences estimation (equation (8)) using cohorts $c \in \{2015, 2017, 2019, 2021\}$ over a bi-annual sample period of 2013–2021. The treatment group consists of workers who first sign an NC in year c. The control group consists of never-treated workers who changed jobs in year c. Standard errors clustered by worker. 95% confidence intervals shown.

Online Appendix A: Proofs Omitted from Main Text

A.1 Social Planner's Solution: Existence, Uniqueness, Comparative Statics

We show that for every parameter set (r, ρ, λ) where $r > \rho$, the surplus function $\mathscr{S}(i) = -\frac{1}{2}i^2 + ri + \frac{e^{-\lambda\Delta i}}{\lambda}$ admits exactly one maximizer $i^\star > 0$. The first derivative of the surplus function is $\mathscr{S}'(i) = r - i - \Delta e^{-\lambda\Delta i}$. At the origin, the derivative is $\mathscr{S}'(0) = r - \Delta = \rho > 0$, implying that zero investment is never optimal. As $i \to \infty$, the exponential term vanishes and $\mathscr{S}'(i) \sim r - i \to -\infty$. Since $\mathscr{S}'(i)$ is continuous and changes sign from positive to negative, it must equal zero at least once for some $i > 0$.

To prove uniqueness, we examine the second derivative, $\mathscr{S}''(i) = -1 + \lambda\Delta^2 e^{-\lambda\Delta i}$. This second derivative is strictly decreasing in i.

- If $\lambda\Delta^2 \leq 1$, then $\mathscr{S}''(i) < 0$ for all $i \geq 0$. This means $\mathscr{S}(i)$ is globally strictly concave, and the critical point identified by the first-order condition is the unique global maximizer.
- If $\lambda\Delta^2 > 1$, then $\mathscr{S}''(i)$ starts non-negative at $i = 0$ and crosses zero exactly once at an inflection point $i_0 > 0$. For all $i > i_0$, the function $\mathscr{S}(i)$ is strictly concave. Because $\mathscr{S}'(0) = \rho > 0$ and $\mathscr{S}'(i)$ is non-decreasing on the interval $[0, i_0]$ (since $\mathscr{S}''(i) \geq 0$ on this interval), any solution to $\mathscr{S}'(i) = 0$ must lie in the region (i_0, ∞) where the function is strictly concave.

Thus, in all cases, there is a single critical point which is the unique global maximizer, $i^\star$.

LEMMA 1. *At the optimal investment level $i^\star$, the surplus function must be locally concave. We prove that the second-order condition is strictly satisfied, e.g., $\mathscr{S}''(i^\star) < 0$.*

Proof. Suppose for contradiction that $\mathscr{S}''(i^\star) = -1 + \lambda\Delta^2 e^{-\lambda\Delta i^\star} \geq 0$. As shown in Appendix A.1, the function $\mathscr{S}''(i)$ is strictly decreasing in i. This implies that for all $i \in [0, i^\star)$, it must be that $\mathscr{S}''(i) > \mathscr{S}''(i^\star) \geq 0$. Therefore, the first derivative $\mathscr{S}'(i)$ must be strictly increasing on the interval $[0, i^\star]$. This leads to a contradiction, since we know $\mathscr{S}'(0) = \rho > 0$ and $\mathscr{S}'(i^\star) = 0$. A function cannot start at a positive value and be strictly increasing to a value of zero. Hence, the premise is false and it must be that $\mathscr{S}''(i^\star) < 0$. ■

We now analyze how the optimal investment and quit probabilities respond to changes in the economic environment.

LEMMA 2 (Comparative Statics of Investment)**.** *The optimal investment $i^\star$ responds to parameter changes as follows: (i) it is strictly increasing in internal productivity, $\partial i^\star/\partial r > 0$; (ii) it is strictly increasing in the outside option parameter, $\partial i^\star/\partial\lambda > 0$; and (iii) its response to external productivity ρ is ambiguous, with $\partial i^\star/\partial\rho \gtrless 0$ if and only if $\lambda\Delta i^\star \lessgtr 1$.*

Proof. Let $\phi := \lambda\Delta i^\star$. Applying the implicit function theorem to Equation 2, we have $\partial i^\star/\partial x = -F_x/F_i$, where $F(i,x) = r - i - \Delta e^{-\lambda\Delta i}$. As shown in Lemma 1, the second-order condition ensures that the denominator $F_i = \mathscr{S}''(i^\star)$ is strictly negative. The signs of the comparative statics are therefore determined by the signs of the partial derivatives $F_x = \partial F/\partial x$:

$$
\begin{aligned}
F_r &= 1 - e^{-\phi}(1-\phi) > 0, \\
F_\lambda &= \Delta^2 i^\star e^{-\phi} > 0, \\
F_\rho &= e^{-\phi}(1-\phi).
\end{aligned}
$$

Since $F_i < 0$, the signs of $\partial i^\star/\partial r$ and $\partial i^\star/\partial\lambda$ are positive. The sign of $\partial i^\star/\partial\rho$ depends on the term $(1-\phi)$, proving the proposition. ■

Higher internal productivity (r) and a less favorable distribution of outside options (higher λ, meaning lower mean $1/\lambda$) both unambiguously increase the incentive to invest. The effect of higher external productivity (ρ) depends on whether quits are common ($\phi < 1$) or rare ($\phi > 1$).

LEMMA 3 (Comparative Statics of the Quit Probability)**.** *The equilibrium quit probability, $q^\star = e^{-\lambda \Delta i^\star}$, is strictly decreasing in internal productivity r and in the outside option parameter λ. The effect of external productivity ρ is ambiguous.*

Proof. Let $\phi := \lambda \Delta i^\star$. The derivative is $dq^\star/dx = -q^\star(d\phi/dx)$, so its sign is opposite to that of $d\phi/dx = d(\lambda \Delta i^\star)/dx$.

- For r: $d\phi/dr = \lambda\Delta(\partial i^\star/\partial r) + \lambda i^\star > 0$. Thus, $dq^\star/dr < 0$.
- For λ: $d\phi/d\lambda = \lambda\Delta(\partial i^\star/\partial \lambda) + \Delta i^\star > 0$. Thus, $dq^\star/d\lambda < 0$.
- For ρ: $d\phi/d\rho = \lambda\Delta(\partial i^\star/\partial \rho) - \lambda i^\star$. The sign is ambiguous as it involves the difference of two terms, with $\partial i^\star/\partial \rho$ itself having an ambiguous sign.

■

A.2 Necessary and Sufficient Conditions for a Non-Zero Wage with an NC

We analyze the conditions under which the firm offers a strictly positive wage, $w_1 > 0$, as part of an optimal contract with an NC. Let $\Sigma_1(w_1)$ denote the joint surplus. Since $\Sigma_1'(0) = 0$, any move from $w_1 = 0$ raises surplus only if $\Sigma_1''(0) > 0$. Equivalently, if we define $H(w_1) \equiv \frac{d\Sigma_1}{dw_1}(w_1)$, then $H(0) = 0$ and local convexity at zero requires $H'(0) > 0$.

Recall that the firm's investment schedule is

$$i_1(w_1) \;=\; r\big(1 - e^{-\lambda w_1}\big),$$

so in particular $i_1(0) = 0$ and hence $\Sigma_1(0) = 0$. The first-order condition for an interior optimum is

$$H(w_1) = (\beta - 1)\big(1 - e^{-\lambda w_1}\big) + \lambda\left(r\, i_1(w_1) - w_1\right) e^{-\lambda w_1},$$

with $H(0) = 0$.

LEMMA 4 (Condition for a Positive Wage with an NC)**.**

(a) A strictly positive equilibrium wage, $w_1^* > 0$, exists if and only if

$$\beta - 2 + \lambda r^2 \;>\; 0, \quad \textit{e.g.,} \quad \beta > 2 - \lambda r^2.$$

(b) A sufficient condition for a positive wage for any $\beta \geq 0$ is

$$\lambda r^2 > 2.$$

Proof. Substitute $i_1(w_1) = r(1 - e^{-\lambda w_1})$ into $H(w_1)$:

$$H(w_1) = (\beta - 1)\big(1 - e^{-\lambda w_1}\big) + \lambda\left(r^2 - r^2 e^{-\lambda w_1} - w_1\right) e^{-\lambda w_1}.$$

Differentiate with respect to w_1:

$$H'(w_1) = \lambda(\beta-1)e^{-\lambda w_1} + \lambda\left(r^2\lambda e^{-\lambda w_1} - 1\right)e^{-\lambda w_1} + \lambda\left(r^2 - r^2e^{-\lambda w_1} - w_1\right)\left(-\lambda e^{-\lambda w_1}\right).$$

Evaluating at $w_1 = 0$:

$$\begin{aligned} H'(0) &= \lambda(\beta-1) + \lambda(\lambda r^2 - 1) + \lambda(r^2 - r^2 - 0)(-\lambda) \\ &= \lambda\beta - \lambda + \lambda^2 r^2 - \lambda \\ &= \lambda\left(\beta - 2 + \lambda r^2\right). \end{aligned}$$

Since $\lambda > 0$, $H'(0) > 0$ if and only if $\beta - 2 + \lambda r^2 > 0$, proving part (a).

For part (b), we require $\beta > 2 - \lambda r^2$ to hold even for the minimum possible value of patience, $\beta = 0$. This gives the condition $0 > 2 - \lambda r^2$, which simplifies to $\lambda r^2 > 2$. This ensures $H'(0) > 0$ for all $\beta \geq 0$, completing the proof. ■

A.3 Proof of Proposition 2

Proof. Let the optimal investments be the roots of the first-order conditions (FOCs) defined as Marginal Benefit - Marginal Cost.

- For the social planner: $F_P(i) \equiv \left(r - \Delta e^{-\lambda\Delta i}\right) - i = 0$
- For the firm: $F_F(i) \equiv \left(r - re^{-\lambda ri}\right) - i = 0$

The firm chooses i_1 such that $F_F(i_1) = 0$. By the second-order condition for a maximum, the derivative of the firm's FOC must be negative at the optimum, so $\partial F_F/\partial i < 0$. This means $F_F(i)$ is a decreasing function of investment i. Because $F_F(i)$ is decreasing, we can determine the sign of the investment distortion by evaluating the sign of $F_F(i^*)$. If $F_F(i^*) < 0$, then $i_1 < i^*$ (under-investment). If $F_F(i^*) > 0$, then $i_1 > i^*$ (over-investment).

We evaluate the firm's FOC at the social optimum i^*:

$$F_F(i^*) = r - re^{-\lambda ri^*} - i^*$$

From the planner's FOC, we know $F_P(i^*) = 0$, which can be rearranged to get $r - i^* = \Delta e^{-\lambda\Delta i^*}$. Substituting this into the expression for $F_F(i^*)$ yields:

$$F_F(i^*) = (r - i^*) - re^{-\lambda ri^*} = \Delta e^{-\lambda\Delta i^*} - re^{-\lambda ri^*}$$

The sign of this expression is given by the sign of $f(\Delta) - f(r)$, where we define the auxiliary function $f(K) = Ke^{-\lambda Ki^*}$. The derivative, $f'(K) = e^{-\lambda Ki^*}(1 - \lambda Ki^*)$, determines if $f(K)$ is increasing or decreasing over the interval $[\Delta, r]$.

A sufficient condition for **under-investment** ($i_1 < i^*$) is $F_F(i^*) < 0$, which requires $f(\Delta) < f(r)$. Since $\Delta < r$, this holds if $f(K)$ is increasing on the interval $[\Delta, r]$. This requires $f'(K) > 0$, for which a sufficient condition is $1 - \lambda ri^* > 0$.

A sufficient condition for **over-investment** ($i_1 > i^*$) is $F_F(i^*) > 0$, which requires $f(\Delta) > f(r)$. Since $\Delta < r$, this holds if $f(K)$ is decreasing on the interval $[\Delta, r]$. This requires $f'(K) < 0$, for which a sufficient condition is $1 - \lambda\Delta i^* < 0$. ■

A.4 Existence of the Boundary Wage

Here, we formally prove the existence of the boundary wage, w_{bound}. This is the wage satisfying the fixed-point condition $w = r\,i_0(w)$, where $i_0(w)$ is the firm's optimal investment response.

LEMMA 5 (Existence and Positivity of the Boundary Wage)**.**

1. *There exists at least one non-negative boundary wage* $w_{\text{bound}} \geq 0$ *solving the equation* $w = r\,i_0(w)$.
2. *A strictly positive solution* $w_{\text{bound}} > 0$ *exists if and only if* $r\,\lambda\,(r-\rho) > 1$.

Proof. We define the function $\Phi(w) = r\,i_0(w)$ and seek a fixed point where $w = \Phi(w)$.

(a) Existence of a non-negative fixed point.

The firm's optimal investment $i_0(w)$ is the unique root of the first-order condition $H(i_0, w) = 0$, where

$$H(i,w) = r\big(1 - e^{-\lambda(w-\rho\,i)}\big) - (r\,i - w)\,\lambda\,\rho\,e^{-\lambda(w-\rho\,i)} - i.$$

The function $H(i,w)$ is twice continuously differentiable. Strict concavity of the firm's profit with respect to i is confirmed by the sign of $\partial H/\partial i$:

$$\frac{\partial H}{\partial i} = -2\lambda\rho r e^{-\lambda(w-\rho i)} - (ri - w)(\lambda\rho)^2 e^{-\lambda(w-\rho i)} - 1 < 0.$$

Given that $H(0,w) > 0$ for $w > 0$ and $H(i,w) \to -\infty$ as $i \to \infty$, the concavity of H in i guarantees a unique positive root $i_0(w)$ for any $w > 0$. The Implicit Function Theorem then ensures that $i_0(w)$ is a continuous function. Hence, $\Phi(w) = r\,i_0(w)$ is also continuous.

To prove the existence of a fixed point, we use Brouwer's Fixed Point Theorem. Let $i_{\text{bench}} \geq 0$ be the unique solution to $i = r(1 - e^{-\lambda(r-\rho)i})$. Define the set $S = [0, W_{\max}]$, where $W_{\max} = r\,i_{\text{bench}}$. This set is non-empty, compact, and convex. Using the properties of $H(i,w)$, for any $w \in S$, the root $i_0(w)$ satisfies $0 \leq i_0(w) \leq i_{\text{bench}}$, which ensures that Φ maps the set S into itself. Since $\Phi : S \to S$ is a continuous function on a compact, convex set, a fixed point $w_{\text{bound}} \in S$ must exist. This establishes part (a).

(b) Condition for a strictly positive fixed point.

We know $w = 0$ is always a fixed point since $i_0(0) = 0$ and thus $\Phi(0) = 0$. A strictly positive fixed point $w_{\text{bound}} > 0$ is guaranteed to exist if there is another fixed point besides zero. Since $W_{\max} = r\,i_{\text{bench}}$ is also a fixed point, a positive solution exists if and only if $W_{\max} > 0$, which requires $i_{\text{bench}} > 0$.

A positive solution $i_{\text{bench}} > 0$ to the equation $i = r(1 - e^{-\lambda(r-\rho)i})$ exists if and only if the slope of the right-hand side is greater than the slope of the left-hand side (which is 1) at the origin $i = 0$. The slope of the right-hand side at $i = 0$ is:

$$\frac{d}{di} r\big(1 - e^{-\lambda(r-\rho)\,i}\big)\bigg|_{i=0} = r\,\lambda\,(r-\rho).$$

Hence, $i_{\text{bench}} > 0$ if and only if $r\,\lambda\,(r-\rho) > 1$. This is the necessary and sufficient condition for the existence of a strictly positive boundary wage, which completes the proof of part (b). ■

LEMMA 6. *For* $\beta = 1$*, at any unconstrained optimal wage* w^*_{unc} *that maximizes the joint surplus:*

1. *The slope of the investment response function is strictly positive:* $\frac{di_0}{dw_0} > 0$.
2. *The optimal wage exceeds the worker's marginal product:* $w^*_{unc} > r i_0(w^*_{unc})$.

The proof for the first point proceeds by contradiction. The second point follows directly from the first.

Proof. The firm's choice of investment i_0 for a given wage w_0 is defined implicitly by the first-order condition of its own profit-maximization problem, which we can write as $G(i_0, w_0) = 0$:

$$G(i_0, w_0) \equiv r\left(1 - e^{-\lambda(w_0 - \rho i_0)}\right) - \lambda\rho(ri_0 - w_0)e^{-\lambda(w_0-\rho i_0)} - i_0 = 0 \tag{12}$$

By the Implicit Function Theorem, the slope of the investment response function, $i_0(w_0)$, is:

$$\frac{di_0}{dw_0} = -\frac{\partial G/\partial w_0}{\partial G/\partial i_0} \tag{13}$$

The denominator, $\partial G/\partial i_0$, is the second derivative of the firm's profit with respect to investment. The second-order condition for a maximum requires $\partial G/\partial i_0 < 0$. Therefore, the sign of the slope is determined by the sign of the numerator:

$$\text{sign}\left(\frac{di_0}{dw_0}\right) = \text{sign}\left(\frac{\partial G}{\partial w_0}\right)$$

Calculating the numerator yields:

$$\frac{\partial G}{\partial w_0} = \lambda e^{-\lambda(w_0-\rho i_0)}\left[r + \rho + \lambda\rho(ri_0 - w_0)\right] \tag{14}$$

Thus, the sign of the slope $\frac{di_0}{dw_0}$ is determined by the sign of the term in the brackets. We now prove by contradiction that an unconstrained optimum cannot exist where $di_0/dw_0 < 0$. An unconstrained optimum w^*_{unc} must satisfy the first-order condition for maximizing the joint surplus. For $\beta = 1$, this condition is:

$$\lambda(ri_0 - w^*_{unc}) + \rho\left(\frac{di_0}{dw_0}\right)\Bigg|_{w^*_{unc}} = 0 \tag{15}$$

Let us hypothesize that an unconstrained optimum w^*_{unc} exists at a point where the slope of the investment curve is negative, e.g., $\frac{di_0}{dw_0} < 0$. Two conditions would have to be met simultaneously at this point.

We can rearrange the first-order condition:

$$\lambda(ri_0 - w^*_{unc}) = -\rho\left(\frac{di_0}{dw_0}\right)$$

Since we hypothesized $\frac{di_0}{dw_0} < 0$ and we know $\rho > 0$, the right-hand side of the equation must be strictly positive. This implies the left-hand side must also be positive. Since $\lambda > 0$, we must have:

$$ri_0 - w^*_{unc} > 0$$

The mathematical condition required for the slope $\frac{di_0}{dw_0}$ to be negative is that the following term is negative:

$$r + \rho + \lambda\rho(ri_0 - w^*_{unc}) < 0$$

We have reached a contradiction. Therefore, our initial hypothesis is false. Any unconstrained optimum must satisfy $\frac{di_0}{dw_0} \geq 0$.

Proof that $w^*_{unc} > ri_0$ This result now follows directly from the first-order wage equation. Rearranging to solve for w^*_{unc} gives:

$$w^*_{unc} = ri_0 + \frac{\rho}{\lambda}\left(\frac{di_0}{dw_0}\right) \tag{16}$$

Therefore, the unconstrained optimal wage is equal to the worker's marginal product plus a strictly positive term, which proves that $w^*_{unc} > ri_0$.

■

A.5 Proof of Proposition 3

Proof. The proof proceeds by establishing an upper bound for the equilibrium investment and showing this bound is below the social optimum. Let the social planner's marginal benefit be $\text{SMB}(i) = r - \Delta e^{-\lambda \Delta i}$, where $\Delta = r - \rho$. The social optimum i^* solves $\text{SMB}(i^*) = i^*$, as previously shown in Equation 2.

1. The Benchmark Case ($w_0 = ri_0$). Consider the boundary case where the firm earns zero profits ex-post. Let the investment at this boundary be i_{bench}. Substituting $w_0 = ri_0$ into the investment FOC yields $i_{\text{bench}} = r(1 - e^{-\lambda \Delta i_{\text{bench}}})$. The firm's private marginal benefit in this case is $\text{PMB}_{\text{bench}}(i) = r(1 - e^{-\lambda \Delta i})$. The difference between the social and private marginal benefit is $\text{SMB}(i) - \text{PMB}_{\text{bench}}(i) = \rho e^{-\lambda \Delta i} > 0$. Since the social marginal benefit is strictly greater than the firm's private marginal benefit for all i, the socially optimal investment level must be strictly greater than the firm's choice, e.g., $i^* > i_{\text{bench}}$.

2. The General Case ($w_0 < ri_0$). In any equilibrium where the firm earns positive profits ex-post, $w_0 < ri_0$. The firm's investment reaction function is strictly increasing in the wage ($\frac{di_0}{dw_0} > 0$). Therefore, any equilibrium investment $i^*_0(w^*_0)$ must be less than the investment level at the boundary benchmark: $i^*_0(w^*_0) < i_0(ri_0) = i_{\text{bench}}$. Combining these two results ($i^*_0 \leq i_{\text{bench}}$ and $i_{\text{bench}} < i^*$), transitivity implies that the equilibrium investment is always strictly below the social optimum:

$$i^*_0 < i^*$$

■

A.6 Proofs of Proposition 4 and its Corollaries

Proof of Proposition 4. We prove this proposition by showing that for any given investment level i and a fixed wage w, the firm's marginal benefit of investment is strictly greater under an NC contract. The marginal benefit (MB) of investment without an NC is:

$$\text{MB}_0(i;w) = \underbrace{r(1 - e^{-\lambda(w - \rho i)})}_{\text{Marginal Product Effect}} - \underbrace{(ri - w)\lambda \rho e^{-\lambda(w - \rho i)}}_{\text{Hold-up Cost}}$$

With an NC, the hold-up cost is eliminated:

$$\text{MB}_1(i;w) = r(1 - e^{-\lambda w})$$

The difference in marginal benefits is:

$$\begin{aligned}
\text{MB}_1(i;w) - \text{MB}_0(i;w) &= r(1 - e^{-\lambda w}) - \left[r(1 - e^{-\lambda(w - \rho i)}) - (ri - w)\lambda \rho e^{-\lambda(w - \rho i)} \right] \\
&= r\left(e^{-\lambda(w - \rho i)} - e^{-\lambda w} \right) + (ri - w)\lambda \rho e^{-\lambda(w - \rho i)} \\
&= \underbrace{re^{-\lambda w}(e^{\lambda \rho i} - 1)}_{\text{Term A}} + \underbrace{(ri - w)\lambda \rho e^{-\lambda(w - \rho i)}}_{\text{Term B}}
\end{aligned}$$

Term A is strictly positive since $\rho > 0$ and $i > 0$. Term B is non-negative for any viable wage, which requires $w \leq ri$. Thus, the difference is strictly positive. As $\text{MB}_1(i;w) > \text{MB}_0(i;w)$ for all i, and both functions are

decreasing in i, the investment level that sets the marginal benefit to its cost must be higher for the NC case. Therefore, $i_1(w) > i_0(w)$.

Part (b): Comparing Equilibrium Investments for $\beta = 1$

When the worker is perfectly patient ($\beta = 1$), the equilibrium wage is set to the worker's marginal product: $w_\delta^* = ri_\delta^*$.

No-NC Equilibrium (i_0^*): Substituting $w_0^* = ri_0^*$ into the investment FOC for the No-NC case, the hold-up cost term disappears, yielding the implicit solution for i_0^*:

$$i_0^* = r(1 - e^{-\lambda(ri_0^* - \rho i_0^*)}) \quad \Longrightarrow \quad i_0^* = r(1 - e^{-\lambda(r-\rho)i_0^*})$$

NC Equilibrium (i_1^*): Substituting $w_1^* = ri_1^*$ into the NC investment function yields:

$$i_1^* = r(1 - e^{-\lambda r i_1^*})$$

Comparison: Consider the function $g(i;K) = r(1 - e^{-\lambda Ki})$. The equilibrium investments i_0^* and i_1^* are the fixed points solving $i^* = g(i^*;K)$ for $K = r - \rho$ and $K = r$, respectively. Since $r > \rho > 0$, we have $r > r - \rho$. The function $g(i;K)$ is strictly increasing in K. Therefore, for any $i > 0$, $g(i;r) > g(i;r-\rho)$. Because the function $g(i;r)$ is strictly greater than $g(i;r-\rho)$, its intersection with the 45-degree line ($y = i$) must occur at a higher value. Thus, $i_1^* > i_0^*$. ■

Proof of Corollary 1. The proofs rely on the main proposition that equilibrium investment is higher with an NC, $i_1 > i_0$. For this proof, we use the benchmark condition that for $\beta = 1$, the equilibrium wage is set at the viability frontier, so $w_1 = ri_1$ and $w_0 = ri_0$.

(a) Wages: The result $w_1 > w_0$ follows directly from $i_1 > i_0$ and the wage rules.

$$w_1 = ri_1 > ri_0 = w_0.$$

(b) Quit Probabilities: The No-NC retention threshold is $T_0 = w_0 - \rho i_0 = (r-\rho)i_0$. The chain of inequalities $w_1 = ri_1 > ri_0 > (r-\rho)i_0 = T_0$ shows the NC retention threshold is strictly higher, which implies the quit probability $q_1 = e^{-\lambda w_1}$ is strictly lower than $q_0 = e^{-\lambda T_0}$.

(c) Bonus Ambiguity: The upfront bonus B depends on the worker's ex-ante expected utility. The difference in utility, $E_0 - E_1$, consists of a negative wage effect ($w_0 - w_1$) and a positive option value effect from the higher quit probability under the No-NC contract. Since these effects have opposing signs, the comparison is ambiguous. ■

Proof of Corollary 2. The upfront bonus is $B_\delta = \mu^0 - \beta E_\delta[U_W]$. To show $B_1 > B_0$, we must show the premise $w_1 < w_0$ implies $E_1[U_W] < E_0[U_W]$. The expected utilities are $E_1[U_W] = w_1 + \frac{1}{\lambda}e^{-\lambda w_1}$ and $E_0[U_W] = w_0 + \frac{1}{\lambda}e^{-\lambda(w_0 - \rho i_0)}$. Let $h(w) = w + \frac{1}{\lambda}e^{-\lambda w}$. This function is strictly increasing in w. The premise $w_1 < w_0$ implies $E_1 = h(w_1) < h(w_0)$. Furthermore, since $\rho i_0 > 0$, we have $-\lambda(w_0 - \rho i_0) > -\lambda w_0$, which implies $E_0 > h(w_0)$. Combining these gives the chain of inequalities $E_1 < h(w_0) < E_0$, proving $E_1[U_W] < E_0[U_W]$. Therefore, the bonus B_1 must be strictly higher than B_0. ■

A.7 Proofs of Proposition 5 and Corollaries

Proof. The proof first addresses the case of a perfectly patient worker ($\beta = 1$), discusses the ambiguity when $\beta \in (0,1)$, and concludes by showing NC is always preferred with $\beta = 0$

Perfectly Patient Worker ($\beta = 1$)

The analysis hinges on the derivative of the No-NC contract's surplus (Σ_0^*) with respect to skill generality (ρ). Let $q_0^*(\rho) = e^{-\lambda(r-\rho)i_0^*(\rho)}$ be the equilibrium quit probability, where $i_0^*(\rho)$ is the equilibrium investment. The total derivative of the surplus is given by the chain rule:

$$\frac{d\Sigma_0^*}{d\rho} = \frac{\partial \Sigma_0}{\partial \rho} + \frac{\partial \Sigma_0}{\partial i_0}\frac{di_0^*}{d\rho}$$

We find the closed-form expression for each component evaluated at the equilibrium.

- The direct effect of ρ on the surplus is: $\frac{\partial \Sigma_0}{\partial \rho} = i_0^* q_0^*$.
- The indirect effect of investment on surplus, using an equilibrium identity, simplifies to: $\frac{\partial \Sigma_0}{\partial i_0} = \rho q_0^*$.
- The change in equilibrium investment with respect to ρ is found by the implicit function theorem on the investment FOC, yielding: $\frac{di_0^*}{d\rho} = -\frac{r\lambda i_0^* q_0^*}{1-\lambda r(r-\rho)q_0^*}$.

Substituting these into the total derivative formula gives:

$$\begin{aligned}\frac{d\Sigma_0^*}{d\rho} &= i_0^* q_0^* + (\rho q_0^*)\left(-\frac{r\lambda i_0^* q_0^*}{1-\lambda r(r-\rho)q_0^*}\right) \\ &= i_0^* q_0^*\left(\frac{1-r\lambda(r-\rho)q_0^* - r\lambda\rho q_0^*}{1-r\lambda(r-\rho)q_0^*}\right) = \frac{i_0^* q_0^*\left(1-\lambda r^2 q_0^*\right)}{1-\lambda r(r-\rho)q_0^*}\end{aligned}$$

The denominator, $1-\lambda r(r-\rho)q_0^*$, is positive. This follows directly from the second-order condition of the firm's profit maximization. For the equilibrium investment i_0^* to be a profit maximum, the second derivative of the profit function with respect to investment must be negative, which is precisely the condition $\lambda r(r-\rho)q_0^* - 1 < 0$. Since the denominator is positive, the sign of the derivative is determined entirely by the sign of the numerator. Let's define this sign-determining function as $f(\rho) = 1-\lambda r^2 q_0^*(\rho)$.

We show that $f(\rho)$ is continuous and strictly decreasing from a positive value at $\rho = 0$ to a negative value as $\rho \to r$, which guarantees a unique root $\hat{\rho}$ where $f(\hat{\rho}) = 0$.

- At $\rho = 0$, we must show $f(0) = 1-\lambda r^2 q_0^*(0) > 0$. This is equivalent to showing $\lambda r^2 q_0^*(0) < 1$. We use the equilibrium condition for investment at $\rho = 0$, which is $i_0^* = r(1-q_0^*(0))$, and the definition of the quit probability, $q_0^*(0) = e^{-\lambda r i_0^*}$. By rearranging the investment condition and substituting, we can show the expression is equivalent to $\frac{\lambda r i_0^*}{e^{\lambda r i_0^*}-1}$. From the fundamental inequality $e^x - 1 > x$ for all $x > 0$, we know the denominator is strictly greater than the numerator, which proves the expression is less than 1. Therefore, $f(0) > 0$.
- As $\rho \to r$, $q_0^* \to 1$ and $f(\rho) \to 1-\lambda r^2 < 0$ by the assumption required for an interior solution without an NC.
- The derivative $f'(\rho) = -\lambda r^2 \frac{dq_0^*}{d\rho}$ is strictly negative because the quit probability q_0^* is an increasing function of ρ.

The Intermediate Value Theorem guarantees a unique tipping point $\hat{\rho}$. For all $\rho \in (0, \hat{\rho}]$, $\frac{d\Sigma_0^*}{d\rho} \geq 0$, which implies $\Sigma_0^*(\rho) > \Sigma_0^*(0) = \Sigma_1^*$. Therefore, the No-NC contract is strictly preferred. For the comparative statics of $\hat{\rho}$, we find its explicit formula:

$$\hat{\rho} = r - \frac{r\ln(\lambda r^2)}{\lambda r^2 - 1}$$

Taking partial derivatives shows that $\frac{\partial \hat{\rho}}{\partial r} > 0$ and $\frac{\partial \hat{\rho}}{\partial \lambda} > 0$.

Ambiguity for the General Case ($\beta < 1$)

For an arbitrary $\beta < 1$, the effect of ρ on the No-NC surplus is ambiguous. The total derivative is $\frac{d\Sigma_0^*}{d\rho} = \frac{\partial \Sigma_0}{\partial \rho} + \frac{\partial \Sigma_0}{\partial i_0}\frac{di_0^*}{d\rho}$. The direct effect, $\partial\Sigma_0/\partial\rho = i_0^* e^{-\lambda T_0^*}[\beta - \lambda(ri_0^* - w_0^*)]$, is of ambiguous sign. The indirect effect is the product of $\partial\Sigma_0/\partial i_0 = \beta\rho e^{-\lambda T_0^*} > 0$ and the investment response $di_0^*/d\rho$, which is also of ambiguous sign. With ambiguity in both the direct and indirect effects, a general result cannot be determined.

NC always preferred with $\beta = 0$

Let the firm's expected profit for a given contract (w,i) and skill-generality parameter $\rho \geq 0$ be denoted by $\Pi(w,i;\rho)$:

$$\Pi(w,i;\rho) = (ri - w)(1 - e^{-\lambda(w-\rho i)}) - \frac{1}{2}i^2$$

The profit function under an NC, Π_1, is identical to this general function evaluated at $\rho = 0$. The No-NC profit, Π_0, corresponds to the case where $\rho > 0$. The firm's problem is to choose a wage w to maximize its profit, anticipating its own investment response, $i^*(w)$. The firm's value function is $V(w;\rho) = \Pi(w, i^*(w;\rho);\rho)$. The maximized equilibrium profits are $V_1^* = \max_w V(w;0)$ and $V_0^* = \max_w V(w;\rho)$. Let w_0^* be the wage that maximizes profits in the No-NC regime, so $V_0^* = V(w_0^*;\rho)$. By the definition of a maximum, $V_1^* \geq V(w_0^*;0)$. The proof is complete if we can show that $V(w_0^*;0) > V_0^*$. We prove this by showing that the difference $V(w_0^*;0) - V_0^*$ is strictly positive.

$$V(w_0^*;0) - V_0^* = \Pi(w_0^*, i^*(w_0^*;0);0) - \Pi(w_0^*, i^*(w_0^*;\rho);\rho)$$

We decompose this difference by adding and subtracting the term $\Pi(w_0^*, i^*(w_0^*;\rho);0)$:

$$\begin{aligned} V(w_0^*;0) - V_0^* = &[\Pi(w_0^*, i^*(w_0^*;0);0) - \Pi(w_0^*, i^*(w_0^*;\rho);0)] \\ &+ [\Pi(w_0^*, i^*(w_0^*;\rho);0) - \Pi(w_0^*, i^*(w_0^*;\rho);\rho)] \end{aligned}$$

We now prove that both bracketed terms are strictly positive.

1. *The Hold-up Gain:* The second bracketed term compares the profit from the *same* contract, $(w_0^*, i^*(w_0^*;\rho))$, under two different legal regimes. Let $i_0^* = i^*(w_0^*;\rho)$.

$$\begin{aligned} \Pi(w_0^*, i_0^*;0) - \Pi(w_0^*, i_0^*;\rho) &= (ri_0^* - w_0^*)\left[(1 - e^{-\lambda w_0^*}) - (1 - e^{-\lambda(w_0^* - \rho i_0^*)})\right] \\ &= (ri_0^* - w_0^*)e^{-\lambda w_0^*}\left[e^{\lambda\rho i_0^*} - 1\right] \end{aligned}$$

   Since $ri_0^* - w_0^* > 0$, and for $\rho > 0$ and $i_0^* > 0$ the term in brackets is positive, this "Hold-up Gain" is strictly positive.

2. *The Investment Gain:* The first bracketed term evaluates the profit under the NC regime ($\rho = 0$) at two different investment levels: the optimal one for that regime, $i_1^* = i^*(w_0^*;0)$, versus the suboptimal one, $i_0^* = i^*(w_0^*;\rho)$.

$$\Pi(w_0^*, i_1^*;0) - \Pi(w_0^*, i_0^*;0)$$

   By definition, for a given wage w_0^*, the investment i_1^* is chosen to uniquely maximize the function $\Pi(w_0^*, i;0)$ with respect to i. As established in the text, the optimal investment response is strictly higher under an NC, so $i_1^* > i_0^*$. Because i_1^* is the unique maximizer, the profit it generates must be strictly greater than the profit generated by any other investment level, including i_0^*. Therefore, this "Investment Gain" is strictly positive.

Since both components of the difference are strictly positive, their sum must be strictly positive. This proves that $V(w_0^*;0) > V_0^*$. The full chain of inequalities is:

$$V_1^* \geq V(w_0^*;0) > V_0^*$$

Thus, the firm's equilibrium profit is strictly higher with an NC contract.

■

Proof of Corollary 3. When a firm's investment makes a worker more productive for a competitor than for the firm itself ($\rho > r$), a standard employment contract (No-NC) becomes untenable. The firm anticipates that the worker will always be poached post-investment and thus refuses to invest, leading to zero surplus. The NC makes the employment relationship viable, as it allows the firm to realize the internal returns on its investment without the certainty of the worker being hired away. ■

Online Appendix B: Simulations of Theoretical Model

Since the model cannot be solved with a closed form solution in most cases, we use simulations to illustrate how NC usage as well as the effects of NC on wage, investment and quit probability, evolve for workers with different levels of productivity r and patience β. We first set the values of parameters $\{r, \rho, \beta, \lambda\}$. For the NC case, we can solve for $\{i_1^*, w_1^*, B_1^*\}$ using Equations 3 and 4 derived in the main text. Since there is no closed form solution for $\{w_0^*, i_0^*\}$, we need to solve for the No NC case iteratively. Given that i_0^* is a function of w_0, we can solve for $i_0^*(w_0)$ for an arbitrary value of w_0 using the FOC for i_0^*. To find the optimal w_0^*, we plug different values of w_0 and $i_0^*(w_0)$ into the expected joint surplus Σ_0 and search for the w_0 that maximizes Σ_0 over a range of possible values. Given $i_0^* < r$ and $w_0^* \leq r i_0^*$, we know that the optimal second period wage w_0^* must be in the range of $[0, r^2]$. Therefore, we discretize $[0, r^2]$ into 20,000 equally spaced points and find the w_0^* that optimizes the expected joint surplus by a brute-force search.

We are mainly interested in $\beta = \{0, 0.8, 1\}$. $\beta = 0.8$ is chosen as an intermediate case that roughly matches some empirical findings on how much households discounts future earnings (Heckman, Lochner, and Todd 2006; Lawrance 1991). For each β, we simulate for low productivity ($r = 1$) and high productivity ($r = 2$) cases. Additionally, we fix $\lambda = 3$ for all simulations. In each simulation, we plot how NC usage ($Sign(\Delta\Sigma^*)$) as well as differentials in investment (Δi^*), wage (Δw^*), bonus (ΔB^*) and quit probability (Δq^*) between NC and No NC evolve with increases in the ratio between external return and internal return $\frac{\rho}{r}$, which measures how industry-specific the investment is.

In Figure B1, we plot the case with fully patient workers ($\beta = 1$). Under this case, since the firm earns zero second period profit ($w = ri$), so the firm's profit is $-c(i) - B$, and the tradeoff between NC and No NC is equivalent to the tradeoff between training costs and bonus. At lower ρ values, since skills are more firm-specific, the firm is encouraged to invest under both NC and No NC. Investment is higher under NC but the difference is small. The bonus is lower under No NC. Workers need to be compensated less upfront under No NC as the skills have high option values upon separation only when NC is not signed. No NC is used as the firm pays a lower bonus and have lower training cost. As skills become more industry-specific (e.g., higher $\frac{\rho}{r}$ ratio), NC encourages a larger increase in investment compared to No NC as the firm invests less without an NC. Since the firm earns zero second period profit, this means the wage premium becomes larger under NC compared to No NC. The bonus has to be higher under No NC compared to NC since the option values under No NC become small and the firm pays a much higher wage under NC. The firm starts shifting to NC as the lower upfront bonus under NC starts to outweigh the higher training costs from more investment under NC.

In Figure B2, we plot the case with completely impatient workers ($\beta = 0$). NC is always used when

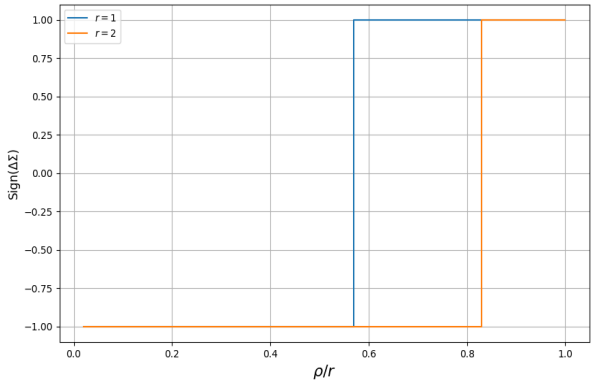
workers completely disregard second period payoff. Since the worker's payoff binds and hinges only on a guaranteed first period wage, the bonus is the same under NC and No NC. Investment is always higher under NC. Second period wage is ambiguous in general. Under the high productivity case, the wage is higher under NC only when the external return is high. In the low productivity case, the wage is always higher under NC. Even though the worker does not care about second period wage, the firm needs to pay the worker a high wage to retain the worker at a higher probability. The wage effect is ambiguous because when the skill is more firm specific (e.g., low $\frac{\rho}{r}$ ratio), the firm still finds it profitable to increase investment under No NC and they may pay more in wages since the outside option is higher under No NC. But as the skills become more general, the firm no longer finds it profitable to invest. Thus, the firm lowers investment and wages paid to worker under No NC and we see a wage premium under NC.

In Figure B3, we plot an intermediate case when workers are partially patient ($\beta = 0.8$). Like before, NC encourages higher investment and the effect is bigger when skills are more transferable. In terms of usage, we see NC is always used in the high productivity case. However, in the low productivity case, we see that NC is used for only the most general skills. Similar to the $\beta = 1$ case, the bonus is higher under NC when skills are more firm-specific. There are several differences from the fully patient worker case. First, the differences in investment between NC and No NC is smaller. This is because worker now wants more front-loaded pay, which forces the firm to pay less in second period wage, which discourages more investment in NC now as well. Second, we see a different wage structure, as the wage is now lower under NC when skills are more firm-specific. Also, bonus is higher under NC for more cases. Again, the difference is due to workers now demanding a more front-loaded wage structure, which raises the bonus under NC.

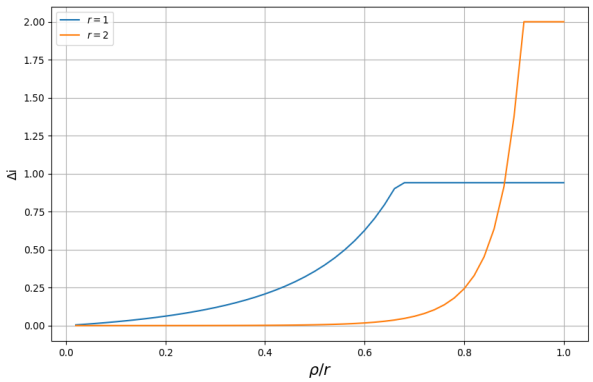
Overall, from all three cases, we see that NC is used more often as skills become more general. This is because the effect of NC encouraging investment is higher as $\frac{\rho}{r}$ becomes larger. The effect of NC on the wage and the bonus, however, is more ambiguous. But as skills become more general, we see a general trend towards higher wage and lower bonus under NC compared to No NC. Lastly, the horizon of the worker, β, plays a critical role here. As worker becomes less patient, NC becomes less effective in encouraging investment as the firm shifts more compensation upfront under NC. However, even though NC no longer facilitates investment in the low patience case, NC is still used.

Figure B1: Simulation of NC Usage, Investment, Wage, Bonus and Quit Probability for Fully Patient Workers ($\beta = 1$)

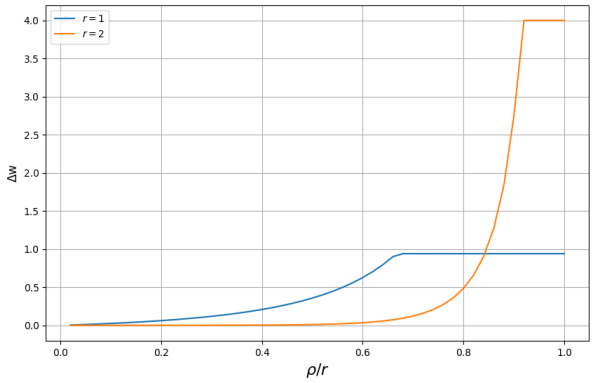
(a) NC Usage = $Sign(\Sigma_1^* - \Sigma_0^*)$



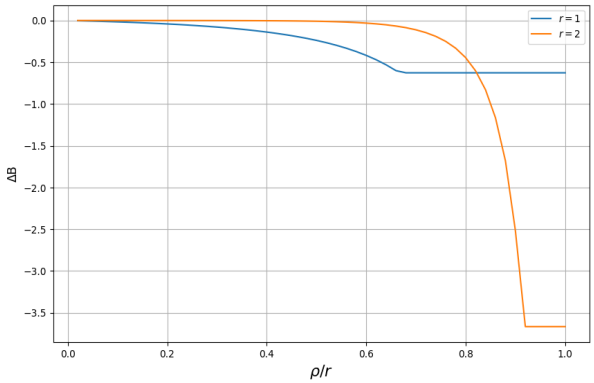
(b) $\Delta_i^* = i_1^* - i_0^*$



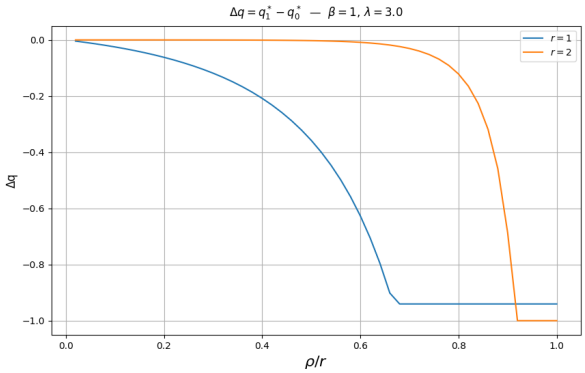
(c) $\Delta_w^* = w_1^* - w_0^*$



(d) $\Delta_B^* = B_1^* - B_0^*$



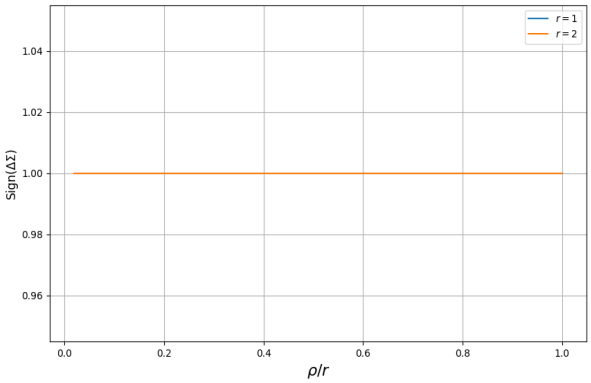
(e) $\Delta_q^* = q_1^* - q_0^*$



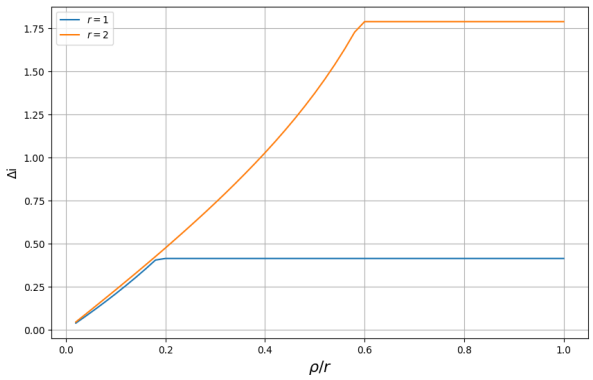
Note: We use simulations to plot how NC usage ($Sign(\Delta\Sigma^*) = Sign(\Sigma_1^* - \Sigma_0^*)$) as well as differentials in investment ($\Delta i^* = i_1^* - i_0^*$), wage ($\Delta w^* = w_1^* - w_0^*$), bonus ($\Delta B^* = B_1^* - B_0^*$) and quit probability ($\Delta q^* = q_1^* - q_0^*$) between NC and No NC evolve with increases in the ratio between external return and internal return $\frac{\rho}{r}$, which measures how industry-specific the investment is. We first set the values of parameters $\{r, \beta, \lambda\}$. In this simulation, we fix $\beta = 1$. We simulate for low productivity ($r = 1$) and high productivity ($r = 2$) cases. Additionally, we fix $\lambda = 3$ for all simulations. Since the NC case has a closed form solution, we plug in $\{r, \beta, \lambda\}$ values and different ρ values to Equation 3 and Equation 4 to solve for $\{w_1^*, i_1^*, B_1^*, q_1^*, \Sigma_1^*\}$. Since the No NC case has no closed form solutions, we need to solve for the No NC case iteratively. Given that i_0^* is a function of w_0, we can solve for $i_0^*(w_0)$ for an arbitrary value of w_0 using Equation 2.4.1. To find the optimal w_0^*, we plug different values of w_0 and $i_0^*(w_0)$ into the expected joint surplus Σ_0 and search for the w_0 that maximizes Σ_0 over a range of possible values. Given $i_0^* < r$ and $w_0^* \leq r i_0^*$, we know that the optimal second period wage w_0^* must be in the range of $[0, r^2]$. Therefore, we discretize $[0, r^2]$ into 20,000 equally spaced points and find the w_0^* that optimizes the expected joint surplus by a brute-force search.

Figure B2: Simulation of NC Usage, Investment, Wage, Bonus and Quit Probability for Fully Impatient Workers ($\beta = 0$)

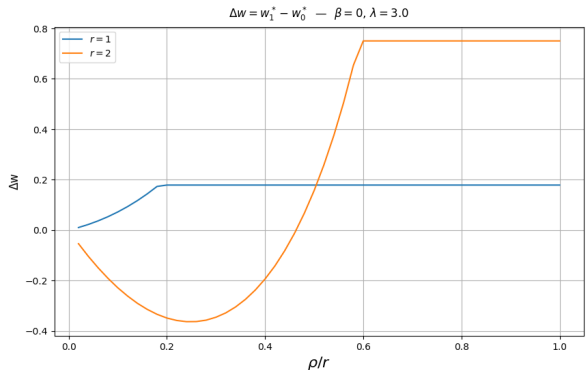
(a) NC Usage = $Sign(\Sigma_1^* - \Sigma_0^*)$



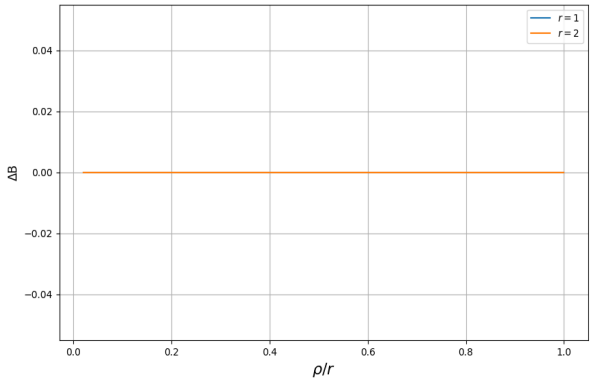
(b) $\Delta_i^* = i_1^* - i_0^*$



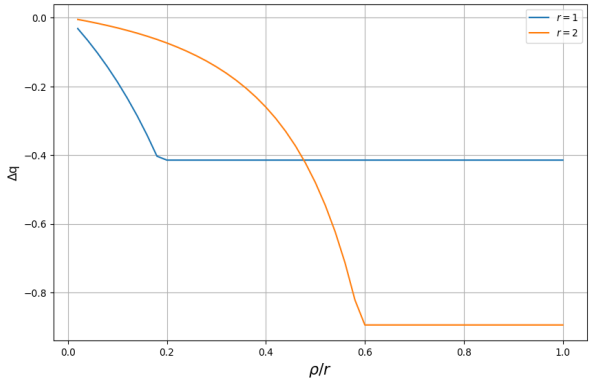
(c) $\Delta_w^* = w_1^* - w_0^*$



(d) $\Delta_B^* = B_1^* - B_0^*$



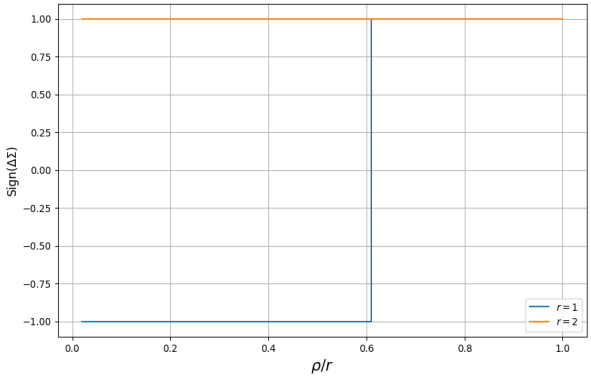
(e) $\Delta_q^* = q_1^* - q_0^*$



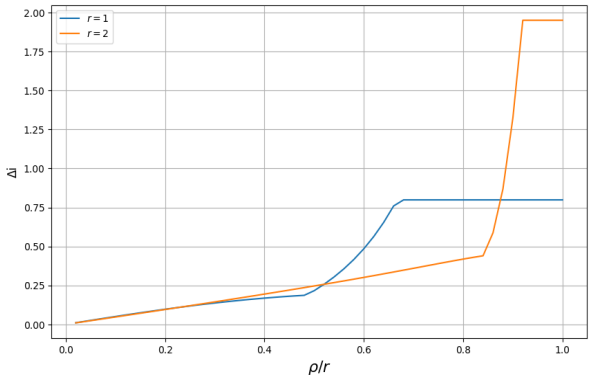
Note: We use simulations to plot how NC usage ($Sign(\Delta\Sigma^*) = Sign(\Sigma_1^* - \Sigma_0^*)$) as well as differentials in investment ($\Delta i^* = i_1^* - i_0^*$), wage ($\Delta w^* = w_1^* - w_0^*$), bonus ($\Delta B^* = B_1^* - B_0^*$) and quit probability ($\Delta q^* = q_1^* - q_0^*$) between NC and No NC evolve with increases in the ratio between external return and internal return $\frac{\rho}{r}$, which measures how industry-specific the investment is. We first set the values of parameters $\{r, \beta, \lambda\}$. In this simulation, we fix $\beta = 0$. We simulate for low productivity ($r = 1$) and high productivity ($r = 2$) cases. Additionally, we fix $\lambda = 3$ for all simulations. Since the NC case has a closed form solution, we plug in $\{r, \beta, \lambda\}$ values and different ρ values to Equation 3 and Equation 4 to solve for $\{w_1^*, i_1^*, B_1^*, q_1^*, \Sigma_1^*\}$. Since the No NC case has no closed form solutions, we need to solve for the No NC case iteratively. Given that i_0^* is a function of w_0, we can solve for $i_0^*(w_0)$ for an arbitrary value of w_0 using Equation 2.4.1. To find the optimal w_0^*, we plug different values of w_0 and $i_0^*(w_0)$ into the expected joint surplus Σ_0 and search for the w_0 that maximizes Σ_0 over a range of possible values. Given $i_0^* < r$ and $w_0^* \leq r i_0^*$, we know that the optimal second period wage w_0^* must be in the range of $[0, r^2]$. Therefore, we discretize $[0, r^2]$ into 20,000 equally spaced points and find the w_0^* that optimizes the expected joint surplus by a brute-force search.

Figure B3: Simulation of NC Usage, Investment, Wage, Bonus and Quit Probability for Partially Patient Workers ($\beta = 0.8$)

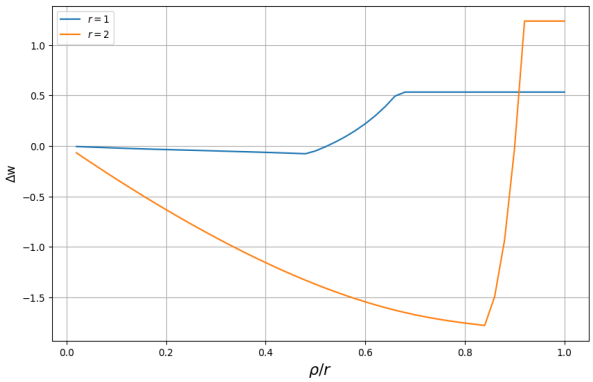
(a) NC Usage = $Sign(\Sigma_1^* - \Sigma_0^*)$



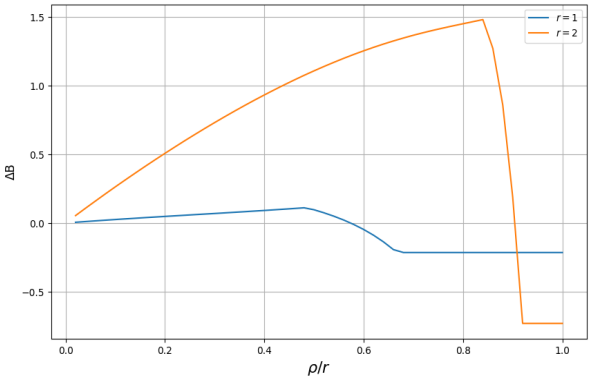
(b) $\Delta_i^* = i_1^* - i_0^*$



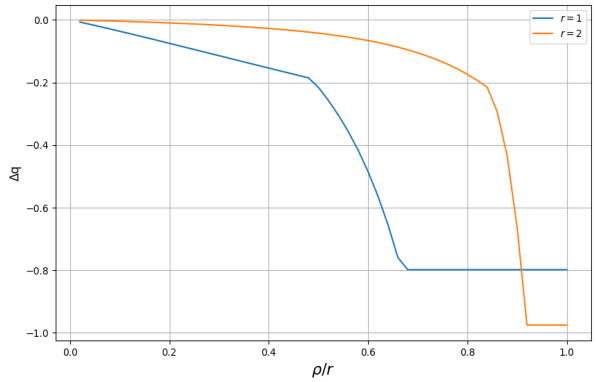
(c) $\Delta_w^* = w_1^* - w_0^*$



(d) $\Delta_B^* = B_1^* - B_0^*$



(e) $\Delta_q^* = q_1^* - q_0^*$



Note: We use simulations to plot how NC usage ($Sign(\Delta\Sigma^*) = Sign(\Sigma_1^* - \Sigma_0^*)$) as well as differentials in investment ($\Delta i^* = i_1^* - i_0^*$), wage ($\Delta w^* = w_1^* - w_0^*$), bonus ($\Delta B^* = B_1^* - B_0^*$) and quit probability ($\Delta q^* = q_1^* - q_0^*$) between NC and No NC evolve with increases in the ratio between external return and internal return $\frac{\rho}{r}$, which measures how industry-specific the investment is. We first set the values of parameters $\{r, \beta, \lambda\}$. In this simulation, we fix $\beta = 0.8$. We simulate for low productivity ($r = 1$) and high productivity ($r = 2$) cases. Additionally, we fix $\lambda = 3$ for all simulations. Since the NC case has a closed form solution, we plug in $\{r, \beta, \lambda\}$ values and different ρ values to Equation 3 and Equation 4 to solve for $\{w_1^*, i_1^*, B_1^*, q_1^*, \Sigma_1^*\}$. Since the No NC case has no closed form solutions, we need to solve for the No NC case iteratively. Given that i_0^* is a function of w_0, we can solve for $i_0^*(w_0)$ for an arbitrary value of w_0 using Equation 2.4.1. To find the optimal w_0^*, we plug different values of w_0 and $i_0^*(w_0)$ into the expected joint surplus Σ_0 and search for the w_0 that maximizes Σ_0 over a range of possible values. Given $i_0^* < r$ and $w_0^* \leq r i_0^*$, we know that the optimal second period wage w_0^* must be in the range of $[0, r^2]$. Therefore, we discretize $[0, r^2]$ into 20,000 equally spaced points and find the w_0^* that optimizes the expected joint surplus by a brute-force search.

Online Appendix C: Theory Extensions: Relaxing the Asymmetric Information Assumption

C.1 Allowing for Renegotiation of the NC Contract

C.1.1 Firm Holds All Bargaining Power

We extend the model to allow for renegotiation after labor demand conditions v are realized. Unlike the baseline model where v is private information, v is observable to both parties, reflecting settings like executive contracts where information is more symmetric. Now the firm can restructure the NC and contractual wages after investment is sunk but before trade or separation occurs. In particular, the firm can match outside offers or allow the worker to leave for an industry competitor in exchange for a buyout payment. If the worker rejects the firm's new offer, the original contract terms still hold. Thus, renegotiation is by mutual consent, as in MacLeod and Malcomson (1993). The timing is as follows:

1. The firm offers an initial contract with an NC $(w_1, B_1, \delta = 1)$ and the worker accepts or rejects.
2. The firm chooses an investment level i_1 at cost $\frac{1}{2}i_1^2$.
3. Labor demand conditions v with continuously differentiable CDF $F()$ are realized and *observed by all parties.*
4. The firm offers a renegotiated contract $\{\bar{w}, \bar{\tau}, \bar{\delta}\}$, where:
   - $\bar{w}$ is the new wage if the parties trade,
   - $\bar{\tau}$ is the termination payment from the worker to firm upon separation,
   - $\bar{\delta} \in \{0,1\}$ is the revised NC status of the contract. If $\bar{\delta} = 0$, the firm waives the NC and the worker's outside option is $v + \rho i_1$; otherwise, it remains v.
5. The worker accepts or rejects the revised offer. If rejected, the original contract governs. The parties then decide whether to trade or separate.

Lemma C.1 (Equilibrium Renegotiation)**.** *In the renegotiation subgame, the firm offers a contract that the worker accepts. This contract implements the efficient action (trade or separation) and ensures the worker obtains flow utility* $\max\{w_1, v\}$.

Proof. After the realization of v, the firm offers a take-it-or-leave-it contract $\{\bar{w}, \bar{\tau}, \bar{\delta}\}$. The worker will accept if the contract yields at least $\max\{w_1, v\}$, which is the best available utility under the original contract where the NC is enforced.

Case 1: Trade is efficient. This occurs when

$$ri_1 \geq v + \rho i_1.$$

The firm prefers to trade. It sets $\bar{\delta} = 1$ to enforce the NC, so the worker's fallback option remains $\max\{w_1, v\}$. The firm offers:

$$\bar{w} = \max\{w_1, v\}, \quad \bar{\delta} = 1.$$

Separation does not occur, so the termination payment $\bar{\tau}$ is irrelevant. The worker accepts and earns utility $\bar{w}$, and the firm earns profit $ri_1 - \bar{w}$.

Case 2: Separation is efficient. This occurs when

$$v+\rho i_1 > r i_1.$$

The firm sets $\bar{\delta}=0$ to waive the NC and allow the worker to quit and access $v+\rho i_1$. To ensure the worker is indifferent between accepting and rejecting, the firm sets:

$$v+\rho i_1-\bar{\tau}=\max\{w_1,v\} \quad \Rightarrow \quad \bar{\tau}=v+\rho i_1-\max\{w_1,v\}.$$

The firm earns $\bar{\tau}$, and the worker earns $\max\{w_1,v\}$. The worker accepts.

Conclusion. In both cases, the firm implements the efficient action and extracts the surplus above $\max\{w_1,v\}$. ■

Proposition C.1 (Renegotiation Ensures Efficient Investments and Allocations)**.** *In the model with renegotiation and industry-specific skills:*

1. *Efficient Turnover: For any investment level i_1, there is trade if and only if $v \le (r-\rho)i_1$.*
2. *Efficient Investment: The firm chooses the efficient investment level i^* that solves*

$$(r-\rho)F\big((r-\rho)i^*\big)+\rho=i^*,$$

which maximizes expected total surplus.

Proof. (1) Efficient Turnover. From Lemma C.1, the firm compares its profit under trade versus separation:

$$\text{Trade: } r i_1-\max\{w_1,v\}, \quad \text{Separation: } v+\rho i_1-\max\{w_1,v\}.$$

The firm prefers trade if and only if:

$$r i_1 \ge v+\rho i_1 \quad \Leftrightarrow \quad v \le (r-\rho)i_1.$$

This implements the socially efficient cutoff rule.

(2) Efficient Investment. The firm chooses i_1 to maximize expected profit:

$$\begin{aligned}\Pi(i_1)=&\int_{-\infty}^{(r-\rho)i_1}\big(r i_1-\max\{w_1,v\}\big)\,dF(v)\\&+\int_{(r-\rho)i_1}^{\infty}\big(v+\rho i_1-\max\{w_1,v\}\big)\,dF(v)-\frac{1}{2}i_1^2.\end{aligned}$$

Let $\Sigma(i_1)$ denote expected total surplus. Then:

$$\Pi(i_1)=\Sigma(i_1)-\mathbb{E}[\max\{w_1,v\}].$$

Since $\max\{w_1,v\}$ does not depend on i_1, the firm maximizes $\Sigma(i_1)$. Differentiating:

$$\frac{d\Sigma}{di_1}=(r-\rho)F((r-\rho)i_1)+\rho-i_1.$$

Setting $\frac{d\Sigma}{di_1}=0$ yields the efficient investment condition:

$$(r-\rho)F((r-\rho)i^*)+\rho=i^*.$$ ■

When renegotiation of the original contract is feasible and the firm can tailor the buyout payment to demand conditions, we show that NCs achieve the first-best. We view this model as more applicable to contracting between firms and chief executive officers, where firms share highly sensitive information to the CEO and both parties may have access to specialized lawyers. Indeed, this model aligns with the empirical findings by Kini, Williams, and Yin (2021), who show that more than 60% of new CEOs have NCs after 2010 and that the CEO is more likely to have an NC if the CEO's skills are easily transferable to other firms in the industry. This model yields a more favorable view of NCs than Shi (2023), where asymmetric information on the entrant's valuation prevents the buyout payment from adjusting to demand conditions. Because the design of the buyout payment is central to the efficiency of the contract, in the next two sections, we consider legal remedies that constrain the firm's choice of buyout payments.

C.1.2 Expectation Damages

We now analyze the game under a different, common legal remedy: expectation damages. In this scenario, the payment from the worker to the firm upon separation is designed to make the firm whole, as if the contract had been fulfilled (Shavell 1984). We model expectation damages by contractually fixing the termination fee at the level of the firm's profit from trade at the original wage, $\bar{\tau} = ri_1 - w_1$. As before, separation requires mutual agreement. We show that the insurance granted to the firm by expectation damages results in over-investment and under-separation.

Proposition C.2 (Inefficient Under-Separation and Over-Investment under Expectation Damages)**.** *When the termination fee is fixed by expectation damages, the equilibrium outcome is inefficient relative to the social planner's benchmark.*

(a) ***Inefficient Stays (Job Lock):*** *Separation occurs less frequently than is socially optimal.*

(b) ***Over-investment:*** *The firm chooses the maximum possible level of investment, $i_1 = r$, which is strictly higher than the socially optimal level i^*.*

Proof. The proof proceeds by first deriving the conditions under which each party agrees to separate, then combining them into a unified separation rule.

Derivation of Separation Conditions For separation to occur, both the firm and the worker must find it in their interest to terminate the contract and exchange the damage payment.

The Firm's Decision. The firm prefers to separate if its payoff from receiving the damage payment exceeds its payoff from continuing the relationship. The actual wage paid if the relationship continues is $\max\{w_1, v\}$.

$$\begin{aligned} \text{Profit from Separation} &> \text{Profit from Continuing} \\ \bar{\tau} &> ri_1 - \max\{w_1, v\} \\ ri_1 - w_1 &> ri_1 - \max\{w_1, v\} \\ w_1 &< \max\{w_1, v\} \end{aligned}$$

This inequality holds if and only if $v > w_1$. Thus, the firm has an incentive to separate only when the worker's outside offer is greater than the contracted wage.

The Worker's Decision. The worker agrees to separate if their utility from moving to an industry competitor is at least as great as their alternative. The worker's value at a competitor is $v+\rho i_1$, from which they must pay the damages $\bar{\tau}$.

$$\begin{aligned}
&\text{Utility from Separation to Competitor} \geq \text{Utility from Alternative} \\
&v+\rho i_1-\bar{\tau} \geq \max\{w_1, v\} \\
&v+\rho i_1-(ri_1-w_1) \geq v \quad (\text{since } v > w_1 \text{ for the firm to agree}) \\
&\rho i_1-ri_1+w_1 \geq 0 \\
&w_1 \geq (r-\rho)i_1
\end{aligned}$$

Thus, the worker agrees to the separation only if the contracted wage is greater than or equal to the productivity gap $(r-\rho)i_1$.

Inefficient Under-Separation The NC is waived and separation to the industry competitor occurs only if *both* derived conditions are met.

- If the contract is structured such that $w_1 < (r-\rho)i_1$, the worker's condition is never met. Separation is impossible.
- If the contract is structured such that $w_1 \geq (r-\rho)i_1$, separation occurs if and only if $v > w_1$.

The socially efficient rule is to separate when $v > (r-\rho)i_1$. In either case, the actual separation cutoff is at or above the efficient cutoff, leading to **under-separation** (job lock). Specifically, inefficient lack of turnover to the industry-competitor occurs when $(r-\rho)i_1 < v < w_1$.

Over-investment The firm, when designing the contract ex-ante, is drawn to the regime where $w_1 \geq (r-\rho)i_1$ because it provides perfect insurance. In this regime, the firm's second-period operating profit is *always* ri_1-w_1:

- If $v \leq w_1$, the worker stays and the firm's profit is ri_1-w_1.
- If $v > w_1$, the worker leaves to the industry competitor and the firm receives a damage payment of $\bar{\tau} = ri_1-w_1$.

Anticipating this perfect insurance, the firm's problem is to choose investment i_1 to maximize its profit:

$$\max_{i_1 \geq 0} \quad \Pi(i_1) = (ri_1-w_1) - \frac{1}{2}i_1^2$$

The first-order condition is $\frac{d\Pi}{di_1} = r-i_1 = 0$, which implies $i_1 = r$. The firm's dominant strategy is to offer a contract satisfying the insurance condition and to choose the maximum investment level, $i_1 = r$. The moral hazard arising from expectation damages results in over-investment relative to the social optimum $i^* < r$. ■

This remedy creates two distinct inefficiencies because it provides perfect insurance to the firm, which distorts both its investment decisions and the joint separation decision. The problem of over-investment arises because the firm has perfect insurance. Because the damage payment $\bar{\tau} = ri_1-w_1$ guarantees the firm its expected profit regardless of whether the worker stays or quits, the firm is fully insulated from the risk of

employee turnover.[25] This creates a moral hazard problem. The firm, when choosing its investment level, no longer needs to discount its returns by the probability of a separation. It invests as if performance is certain, because from its perspective, the return is certain. Since the firm is guaranteed a marginal return of r on its investment, it invests until the marginal cost equals this return, leading to the socially excessive level $i_1 = r$.

This remedy also generates inefficient under-separation because the conditions for mutual consent to separate are stricter than the condition for social efficiency. For a separation to occur, the worker must find it privately optimal to pay the expectation damages, $\bar{\tau} = ri_1 - w_1$. The worker's utility from moving to the industry competitor is $v + \rho i_1 - \bar{\tau}$, which must exceed their reservation utility of v (given that the firm only consents to waive the NC when $v > w_1$). This requirement simplifies to the condition $w_1 \geq (r-\rho)i_1$. Consequently, a socially efficient separation, which requires only that $v > (r-\rho)i_1$, is blocked whenever the worker's private condition ($w_1 \geq (r-\rho)i_1$) or the firm's private condition ($v > w_1$) is not met. This misalignment leads to inefficiently low turnover to the industry competitor.

C.1.3 Training Repayment Programs

We now analyze a modification to the renegotiation game to incorporate a Training Repayment Agreement Program (TRAP), as studied in, for example, Feess and Muehlheusser (2003). In this framework, if the parties mutually agree to waive the NC, the worker must pay the firm a contractually fixed termination fee equal to the firm's investment cost, $\bar{\tau} = \frac{1}{2}i_1^2$. This provision represents a specific application of liquidated damages, a contractual mechanism in which the parties pre-commit to a fixed monetary transfer that is payable upon a specified event, such as breach or early termination (e.g., Spier and Whinston 1995). In this case, the worker pays a pre-determined amount to be released from the NC, thereby compensating the firm for its upfront investment.

Proposition C.3 (Inefficient Matching and Investment under TRAPs)**.** *Relative to the social planner's benchmark, Training Repayment Agreement Programs (TRAPs) create inefficiently low separation and an inefficient level of investment.*

1. ***Under-separation:*** *TRAPs lead to under-separation. In the regime where separation is possible ($i_1 \leq 2\rho$), the threshold for separation is higher than the socially efficient threshold. In the regime where investment is high ($i_1 > 2\rho$), socially efficient separations are completely foreclosed as the worker will never agree to pay the fee.*

2. ***Inefficient Investment:*** *TRAPs lead to an inefficient level of investment. The firm's chosen investment will either be unambiguously excessive (if it chooses a high-investment strategy to prevent separation) or will be pinned down by the need to compensate the worker sufficiently to accept the contract, rather than by maximizing joint surplus.*

Proof. The proof compares the equilibrium outcomes under the TRAP with the social planner's benchmark, where separation is efficient (conditional on investment) if $v > v_{eff} = (r-\rho)i_1$ and investment i^* is chosen to maximize expected total surplus.

1. Under-separation. The proof of under-separation first requires establishing the conditions under which separation is possible. Separation requires both the firm's and the worker's participation constraints to be satisfied. The worker agrees to the TRAP's separation terms only if their utility from doing so is greater than or equal to their utility from rejecting the offer. Formally, for $v > w_1$:

$$v + \rho i_1 - \frac{1}{2}i_1^2 \geq v \quad \Rightarrow \quad \rho i_1 \geq \frac{1}{2}i_1^2 \quad \Rightarrow \quad i_1 \leq 2\rho$$

[25] Technically, the worker has the option to quit and join the non-industry competitor without paying the fee. However, we show formally in the proof that quits to the non-industry competitor do not occur in equilibrium in this model.

This derivation shows that separation is only possible if the firm's investment is within the regime $i_1 \leq 2\rho$. If $i_1 > 2\rho$, the worker will always reject the separation offer.

Next, we consider the firm's incentive constraint, which defines its separation cutoff, $v_c = ri_1 - \frac{1}{2}i_1^2$. For the cases where $i_1 \leq 2\rho$, we compare this firm cutoff to the planner's efficient cutoff, $v_{eff} = (r-\rho)i_1$. The difference is:

$$v_c - v_{eff} = \left(ri_1 - \frac{1}{2}i_1^2\right) - (r-\rho)i_1 = i_1\left(\rho - \frac{i_1}{2}\right)$$

Since we are in the regime where $i_1 \leq 2\rho$ (which implies $\rho \geq i_1/2$), the term $(\rho - i_1/2)$ is non-negative. Thus, $v_c \geq v_{eff}$. The firm requires a higher outside offer v to release a worker than is socially optimal, leading to under-separation.

2. Inefficient Investment. The firm chooses i_1 to maximize its own profit by anticipating the outcomes of its two main strategies.

Strategy A (High Investment): Choose $i_1 > 2\rho$. This makes separation impossible. The firm maximizes $\Pi(i_1) = \mathbb{E}[ri_1 - \max\{w_1, v\}] - \frac{1}{2}i_1^2$, which yields the solution $i_1 = r$.

Strategy B (Low Investment): Choose an investment i_1 in the range $[0, 2\rho]$. In this regime, the firm's expected profit is $\Pi(i_1) = \int_{-\infty}^{v_c}(ri_1 - v)\,dF(v) + \int_{v_c}^{\infty}(\frac{1}{2}i_1^2)\,dF(v) - \frac{1}{2}i_1^2$. The first-order condition is:

$$\frac{d\Pi}{di_1} = (r-i_1)F\left(ri_1 - \frac{1}{2}i_1^2\right)$$

For any investment level $i_1 < r$, this derivative is non-negative. Because the profit function is monotonically increasing with respect to investment in this regime, the profit is maximized at the boundary of the interval, which is precisely $i_1 = 2\rho$.

The firm's final choice is to compare profits under each strategy. Neither $i_1 = r$ nor $i_1 = 2\rho$ is chosen to satisfy the social planner's first-order condition, so investment is always inefficient. If the firm adopts Strategy A ($i_1 = r$), there is no separation in equilibrium and the result is unambiguously over-investment: $i^* < r$. If the firm adopts Strategy B ($i_1 = 2\rho$), the investment level is inefficient, but it could be higher or lower than the social optimum (i^*). ■

TRAPs create inefficiencies because the fixed repayment fee forces the firm into a strategic choice between two distinct investment regimes, neither of which is socially optimal. First, the firm can choose a high-investment strategy. By investing heavily, the firm makes the repayment fee ($\frac{1}{2}i_1^2$) prohibitively expensive for the worker. The worker is effectively locked into the firm or the non-industry outside option; separation to an industry competitor never occurs. Anticipating this lock-in, the firm is incentivized to over-invest relative to the social optimum. Alternatively, the firm can pursue a low-investment strategy, keeping the repayment fee low enough that the worker may be willing to pay it in some cases. However, even in this regime, the outcome is inefficient. The conditions required for both the firm and the worker to mutually agree to the separation are stricter than the condition for social efficiency, which results in less turnover than is socially optimal.

C.2 Allowing for Renegotiation of the No-NC Contract

Proposition C.4. *In a model without a non-compete agreement and renegotiation:*

1. *Efficient Turnover: The worker is retained if and only if it is socially efficient to do so.*
2. *Under-investment: The firm's equilibrium investment level is strictly less than the socially optimal level ($i_0^* < i^*$).*

Proof. The proof is organized in two parts. First, we establish the efficiency of the separation decision. Second, we prove that the firm systematically under-invests.

Part 1: Proof of Efficient Turnover A social planner would retain the worker if and only if the surplus from staying, ri_0, is at least as large as the surplus from separating, $v + \rho i_0$. The efficient retention rule is therefore $v \leq (r - \rho)i_0$.

We show that private incentives align with this rule. After investment i_0 is sunk and the outside offer v is realized, the worker's reservation utility becomes their best available alternative. To prevent the worker from leaving, the firm must offer a wage at least as high as this reservation utility, which is $\max\{w_0, v + \rho i_0\}$.

Since the firm has all bargaining power, it will offer exactly this wage to retain the worker, but only if it is profitable to do so. The firm's retention decision is thus governed by whether output covers the new wage:

$$ri_0 \geq \max\{w_0, v + \rho i_0\}.$$

Assuming the initial contract is viable ($w_0 \leq ri_0$), this single condition for profitable retention becomes equivalent to the social efficiency condition, $ri_0 \geq v + \rho i_0$. If retention is efficient, the firm will profitably retain the worker. If retention is inefficient, the firm cannot make a profitable offer, and the worker separates. Thus, turnover is socially efficient.

Part 2: Proof of Under-investment We prove under-investment by comparing the marginal return on investment for the firm with that of a social planner. The social planner's marginal return is r if the match continues and ρ if it dissolves.

Now consider the firm's private marginal return on investment.

- When the worker separates (which occurs when $v > (r - \rho)i$), the firm's profit is zero, and its marginal return on investment is also zero. This is strictly less than the planner's marginal return of ρ.
- When the worker is retained (which occurs when $v \leq (r - \rho)i$), the firm's profit is $ri - \max\{w_0, v + \rho i\}$. The firm's marginal return is therefore either r (if no wage renegotiation is needed) or $r - \rho$ (if the wage must be increased to match the outside option). In all cases of retention, the firm's marginal return is less than or equal to the planner's marginal return of r.

This proof demonstrates that in every possible outcome, the firm's marginal return on investment is less than or equal to the social planner's, and in many outcomes (all separations and all retentions that require renegotiation), it is strictly less. Because the firm fails to internalize the full social value of its investment—particularly the skill-generality parameter ρ that either benefits a competitor or must be paid back to the worker in a higher wage—its total ex-ante Private Marginal Benefit is strictly lower than the Social Marginal Benefit. Therefore, the firm under-invests relative to the social optimum, $i_0^* < i^*$. ■

Online Appendix D: NCs and Returns to Experience

We estimate the return on experience separately by occupation and industry using Equation 7 to show evidence of NCs encouraging firm-provided general skill training by showing that the return to experience is higher in occupations and industries with higher usage of NC. Estimating Equation 7 by OLS would produce a biased result as θ_i, μ_{ij} may be correlated with both job tenure and NC status. We also cannot directly apply a first-difference estimation using within job wage changes over time as tenure and experience are collinear

within the same job. So we apply the two-step estimation method used by Topel (1991).

$$\frac{log(w_{ij,t}) - log(w_{ij,t-2})}{2} = \beta_2 + \beta_3 + e_{ijt} \tag{17}$$

In the first step, we apply a first-difference estimation only on job stayers, defined as workers who did not switch job between 2017 and 2019 or between 2019 and 2021. Since experience and tenure increase both by 2 years for job stayers, the equation can be reduced to Equation 17. This equation allows us to consistently estimate $\beta_2 + \beta_3$.

$$log(w_{ijt}) - (\widehat{\beta_2 + \beta_3}) \cdot T_{ijt} = \beta_1 + \beta_2 X^0_{ijt} + u_{ijt} \tag{18}$$

To separately identify these parameters, in the second step, we decompose X_{ijt} into $X^0_{ijt} + T_{ijt}$, where X^0_{ijt} denotes the pre-period job experience. Then we can plug in the estimates $\widehat{\beta_2 + \beta_3}$ to Equation 7 and rearrange it to get Equation 18. We estimate Equation 18 with a pooled OLS on the NLSY 2017-2021 sample, when the survey starts collecting information on NC status. This procedure yields estimates for the return to experience, $\hat{\beta}_2$.

Online Appendix E: Tables and Figures

Table E1: Variables Dictionary

Variable	Definition
Job Mobility	
Tenure (Yrs)	Time in years working at a job
1(Main Job Separation btwn 2017 and 2019)	Worker transitions to a new main job, becomes unemployed, or exits the labor force
1(Main Job Mobility btwn 2017 and 2019)	Worker transitions to a new main job
1(Within-Industry Job Mobility btwn 2017 and 2019)	Worker transitions to a new main job within the same industry
Wages and Wage Growth	
Log(Starting Wage)	Log of starting wage for 2017 main job
Log(Wage in 2017)	Log of wage in 2017 for main job
$Log(Wage_{2017}) - Log(Wage_{2015})$	Difference between log of wage in 2017 and log of wage in 2015 for 2017 main job
$Log(Wage_{2019}) - Log(Wage_{2017})$	Difference between log of wage in 2019 and log of wage in 2017 for 2017 main job
Demographics	
Age	Computed as the survey year minus birth year
1(Male)	The respondent is male
1(High School Degree or Higher)	The respondent has attended at least 12 years of school
1(Bachelors Degree or Higher)	The respondent has attended at least 16 years of school
ASVAB Percentile	Percentile achieved on ASVAB test
1(Black)	The respondent is Black
1(Hispanic)	The respondent is Hispanic
Wage Bargaining and Negotiation	
1(Possible to Keep Previous Job)	The respondent was able to keep previous job when offered their main job
1(Negotiate Job Offer)	The respondent negotiated their main job offer
Training	
1(Received Some Training)	Received training in a survey year
1(Received Training Run by Employer)	Received training ran by employer in a survey year
1(Received On-Site Training by Non-Employer)	Received training on-site by non-employer in a survey year
1(Employer Paid for Training)	Employer paid for training in a survey year
1(Employer Paid for Mandatory Training)	Employer paid for mandatory training in a survey year
1(Employer Paid for Voluntary Training)	Employer paid for voluntary training in a survey year
Job Tasks	
1(Use Math Skills Frequently)	Respondent claims to use math at least once a week at main job
1(Supervise Frequently)	Respondent claims to supervise more than half the time at main job
1(Problem Solve Frequently)	Respondent claims to problem solve at least once a week at main job
Other Firm Characteristics	
1(Dislike Job)	Respondent claim to 'Dislike it somewhat' or 'Dislike it very much' when asked about main job
1(Unionized Worker)	Respondent's contract was negotiated by a union or employee association for main job
Firm Size	Number of employees at respondent's main job

Table E2: Confidence in NC Status by Industry

Industry	NC Confidence			Total	Share Very Confident	Share NC Usage
	Very Confident	Somewhat Confident	Not Confident			
AGRICULTURE, FORESTRY AND FISHERIES	33	0	1	34	0.97	0.07
CONSTRUCTION	293	14	4	311	0.94	0.11
OTHER SERVICES	152	9	2	163	0.93	0.13
TRANSPORTATION AND WAREHOUSING	211	12	7	230	0.92	0.14
EDUCATIONAL, HEALTH, AND SOCIAL SERVICES	1169	93	8	1270	0.92	0.08
ACS SPECIAL CODES	191	16	1	208	0.92	0.20
UTILITIES	29	3	0	32	0.91	0.11
INFORMATION AND COMMUNICATION	86	8	0	94	0.91	0.24
FINANCE, INSURANCE, AND REAL ESTATE	312	29	3	344	0.91	0.19
ENTERTAINMENT, ACCOMODATIONS, AND FOOD SERVICES	422	38	3	463	0.91	0.07
PUBLIC ADMINISTRATION	232	21	1	254	0.91	0.08
MANUFACTURING	412	42	4	458	0.90	0.18
MINING	24	3	0	27	0.89	0.24
WHOLESALE TRADE	104	12	1	117	0.89	0.27
RETAIL TRADE	458	49	5	512	0.89	0.15
PROFESSIONAL AND RELATED SERVICES	560	64	7	631	0.89	0.28
TOTAL	4688	413	47	5148	0.91	0.15

Note:
The sample consists of NLSY97 respondents who report NC, NC confidence, and industry status in their 2017 main job. Rows are organized by share 'Very Confident' in response to the NC confidence question. Active duty military respondents are dropped.

Table E3: Relationship Between NCs and Labor Market Outcomes: 2017 Cross-Section

Panel 1: Wages and Wage Growth

Dependent Variables:	Log(Wage)			Wage Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
1(NC)	0.211***	0.144***	0.081***	0.009	0.007	0.004
	(0.027)	(0.022)	(0.024)	(0.014)	(0.014)	(0.018)
Controls	None	Basic	Advanced	None	Basic	Advanced
Weighted Dependent Variable Mean	3.04	3.04	3.04	0.088	0.088	0.088
Fit statistics						
Observations	4,968	4,836	3,141	4,968	4,836	3,141
R^2	0.017	0.296	0.456	0.0001	0.0005	0.026

Panel 2: Training

Dependent Variables:	1(Any Training)			1(Emp Paid for Training)		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
1(NC)	-0.019	-0.025*	-0.017	-0.018*	-0.026**	-0.022
	(0.013)	(0.013)	(0.017)	(0.010)	(0.011)	(0.014)
Controls	None	Basic	Advanced	None	Basic	Advanced
Weighted Dependent Variable Mean	0.112	0.112	0.112	0.071	0.071	0.071
Fit statistics						
Observations	4,968	4,836	3,141	4,968	4,836	3,141
R^2	0.0005	0.007	0.048	0.0006	0.012	0.063

Panel 3: Job Mobility

Dependent Variables:	Tenure (Yrs)			1(Job Mobility Between 2017-2019)		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
1(NC)	0.125	0.002	-0.267	-0.036*	-0.026	-0.032
	(0.197)	(0.198)	(0.234)	(0.020)	(0.020)	(0.024)
Controls	None	Basic	Advanced	None	Basic	Advanced
Weighted Dependent Variable Mean	5.08	5.08	5.08	0.304	0.304	0.304
Fit statistics						
Observations	4,884	4,836	3,141	4,968	4,917	3,177
R^2	0.000	0.017	0.083	0.0008	0.011	0.039

Notes: Output reports β_1 from regression $Y_i = \beta_0 + \beta_1 * NC_i + \beta_2 * X_i + \varepsilon_i$. Standard errors are heteroskedasticity-robust. Basic controls include sex, education, tenure, and potential experience. Advanced controls further add industry and occupation fixed effects, ASVAB percentile, and firm size. Tenure controls are not included in the job mobility panel. All regressions are weighted so as to be nationally representative. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table E4: Relationship Between NCs and Labor Market Outcomes: 2019 Cross-Section

Panel 1: Wages and Wage Growth

Dependent Variables:	Log(Wage)			Wage Growth		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
1(NC)	0.305***	0.220***	0.145***	0.005	0.006	0.028
	(0.051)	(0.044)	(0.052)	(0.033)	(0.033)	(0.043)
Controls	None	Basic	Advanced	None	Basic	Advanced
Weighted Dependent Variable Mean	3.14	3.14	3.14	0.111	0.111	0.111
Fit statistics						
Observations	1,638	1,585	762	1,638	1,585	762
R^2	0.034	0.302	0.574	0.000	0.004	0.060

Panel 2: Training

Dependent Variables:	1(Any Training)			1(Emp Paid for Training)		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
1(NC)	0.011	-0.002	-0.028	0.015	-0.002	-0.031
	(0.029)	(0.029)	(0.043)	(0.024)	(0.024)	(0.038)
Controls	None	Basic	Advanced	None	Basic	Advanced
Weighted Dependent Variable Mean	0.120	0.120	0.120	0.078	0.078	0.078
Fit statistics						
Observations	1,638	1,585	762	1,638	1,585	762
R^2	0.0001	0.007	0.108	0.0004	0.020	0.139

Panel 3: Job Mobility

Dependent Variables:	Tenure (Yrs)			1(Job Mobility Between 2019-2021)		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
1(NC)	0.256	0.237	0.116	-0.051	-0.050	-0.064
	(0.184)	(0.187)	(0.226)	(0.036)	(0.037)	(0.052)
Controls	None	Basic	Advanced	None	Basic	Advanced
Weighted Dependent Variable Mean	5.64	5.64	5.64	0.249	0.249	0.249
Fit statistics						
Observations	1,616	1,585	762	1,638	1,607	771
R^2	0.003	0.008	0.098	0.001	0.010	0.068

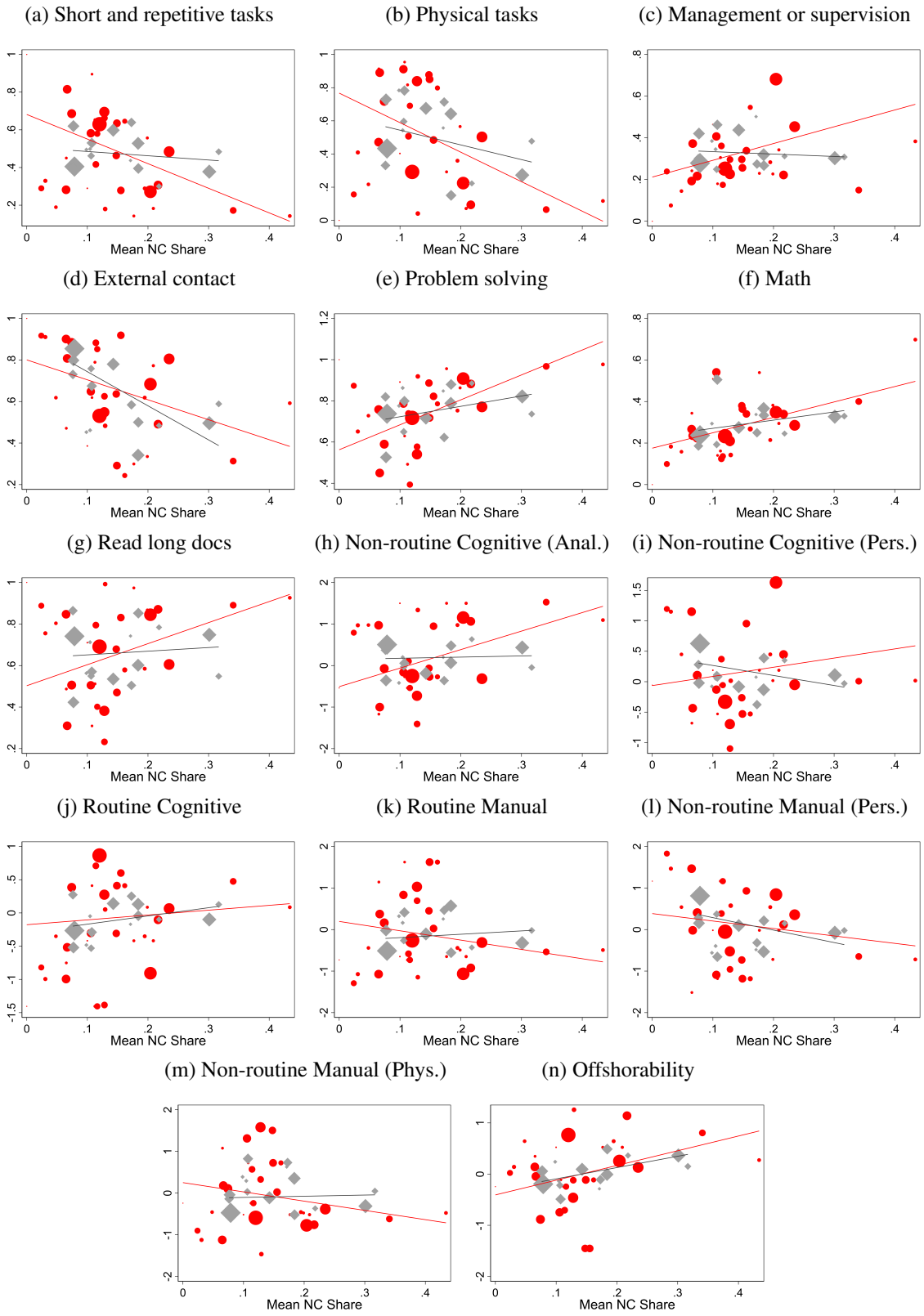
Notes: Output reports β_1 from regression $Y_i = \beta_0 + \beta_1 * NC_i + \beta_2 * X_i + \varepsilon_i$. Standard errors are heteroskedasticity-robust. Basic controls include sex, education, tenure, and potential experience. Advanced controls further add industry and occupation fixed effects, ASVAB percentile, and firm size. Tenure controls are not included in the job mobility panel. All regressions are weighted so as to be nationally representative. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Table E5: The Relationship Between Formal Training and Wages

	Log hourly wage			Log hourly wage growth		
	(1)	(2)	(3)	(4)	(5)	(6)
Training indicator	0.100***	0.0603***	0.0380***	0.00378	-0.000217	0.00309
	(0.0179)	(0.0150)	(0.0136)	(0.00861)	(0.00873)	(0.00892)
Worker characteristics		yes	yes		yes	yes
Ind. and Occ. fixed effects			yes			yes
Observations	18,464	18,464	18,464	17,079	17,079	17,079
R-squared	0.003	0.310	0.438	0.000	0.012	0.018

Notes: Standard errors are clustered at the individual level. The sample restricts to individuals in years 2013–2021 with real wages between 3 and 200. Worker characteristics include tenure, experience and fixed effects for sex, race, age, highest level of education, and ASVAB percentiles. All regressions are weighted so as to be nationally representative. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

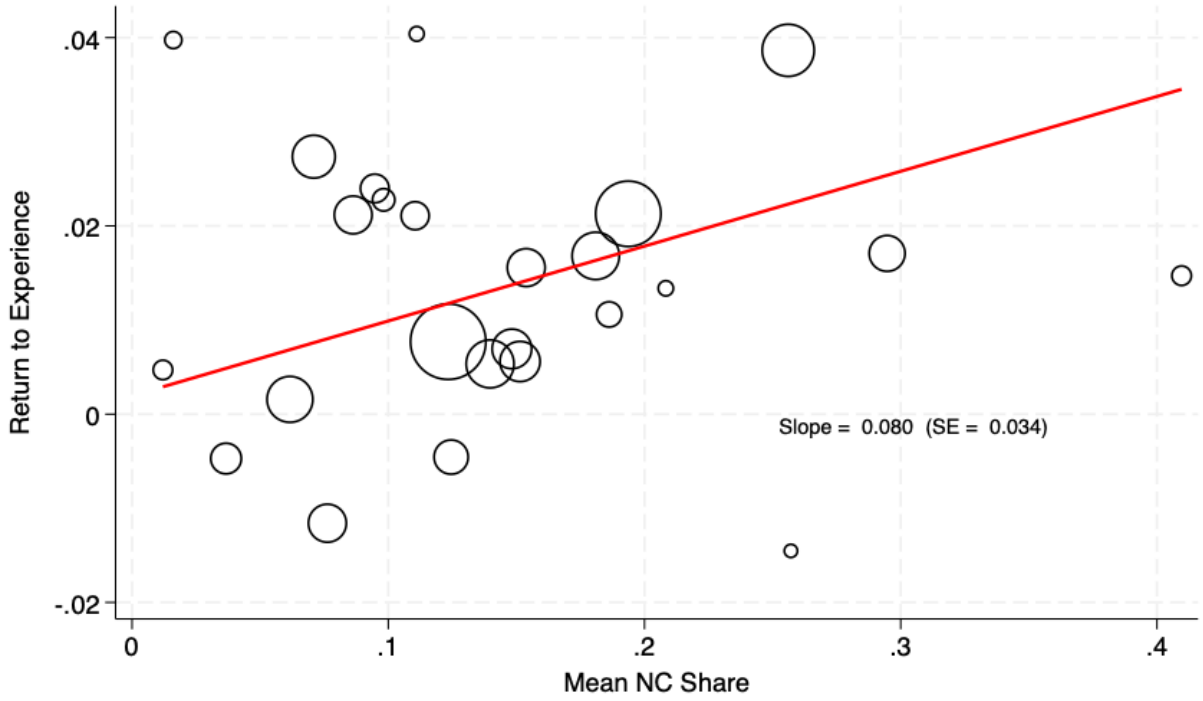
Figure E1: NC Usage and Job Task Content among Industries and Occupations



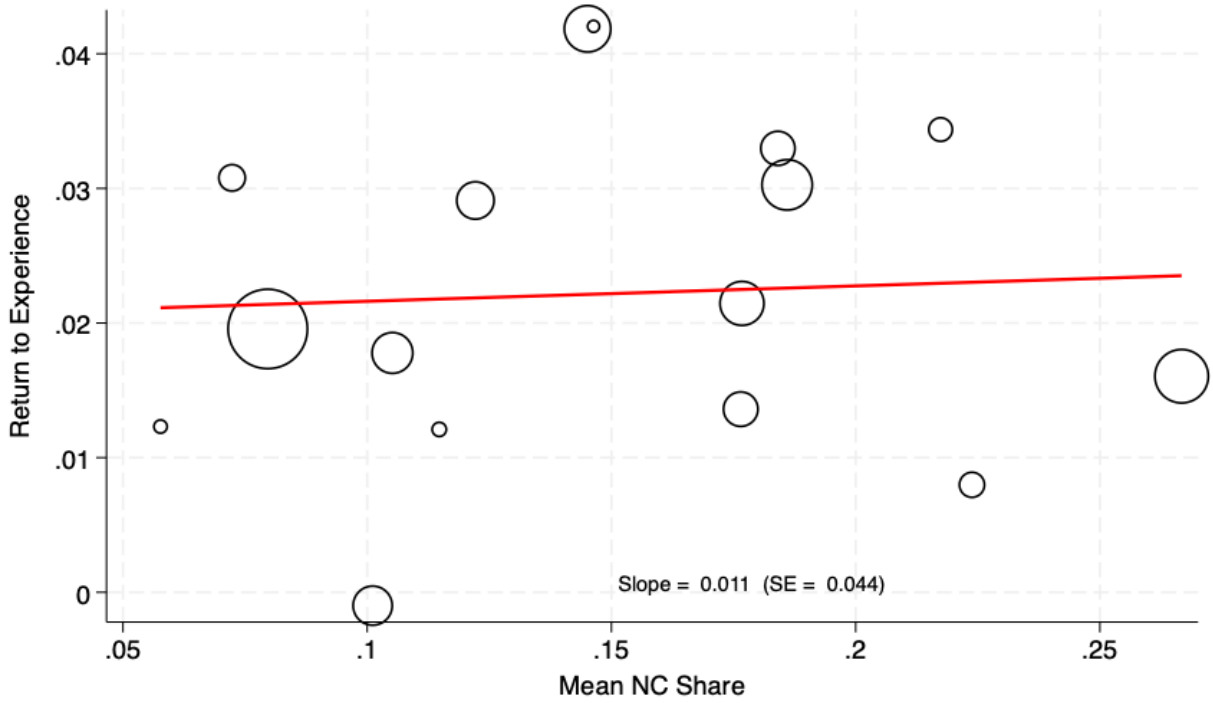
Notes: Gray markers refer to industries and red markers refer to occupations. The analysis plots the average share of workers who are NC signers against the average task content of jobs in each industry/occupation. Sample pools the bi-annual 2017–2021 cross sections of the NLSY97. The routine, non-routine, and offshorability measures are merged from O*NET. Specifically, we look at intensity of (i) non-routine cognitive (analytical), (ii) non-routine cognitive (interpersonal), (ii) routine cognitive, (iv) routine manual, (v) non-routine manual (interpersonal), (vi) non-routine manual (physical), and (vii) offshorability.

Figure E2: NC Usage and Return to Experience

(a) By Occupation

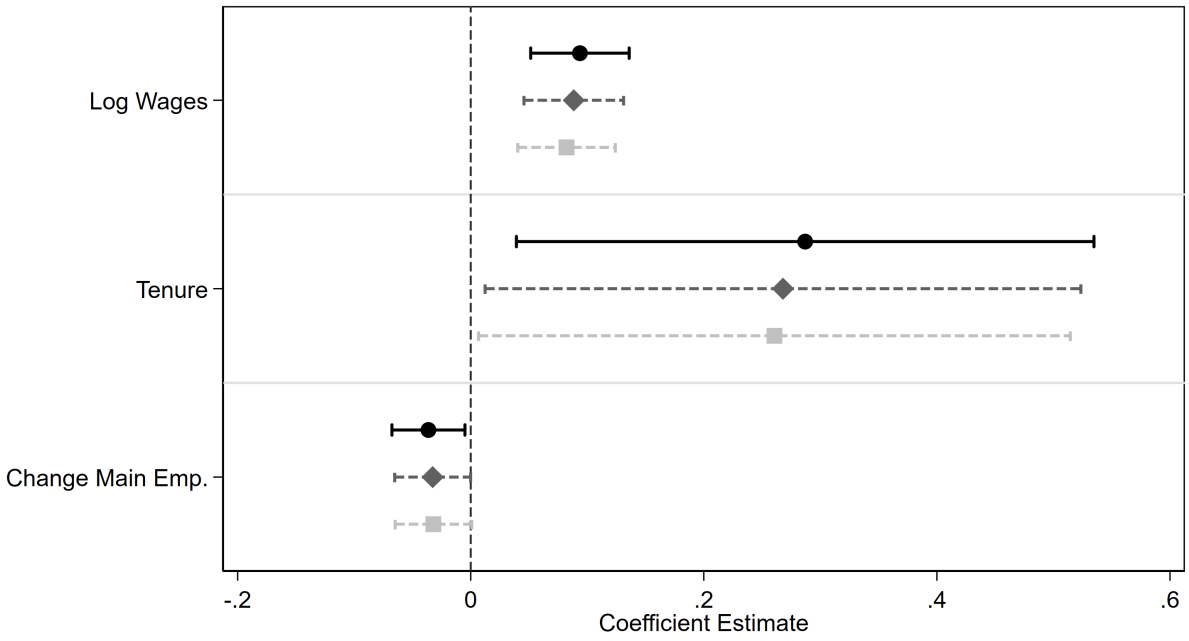


(b) By Industry



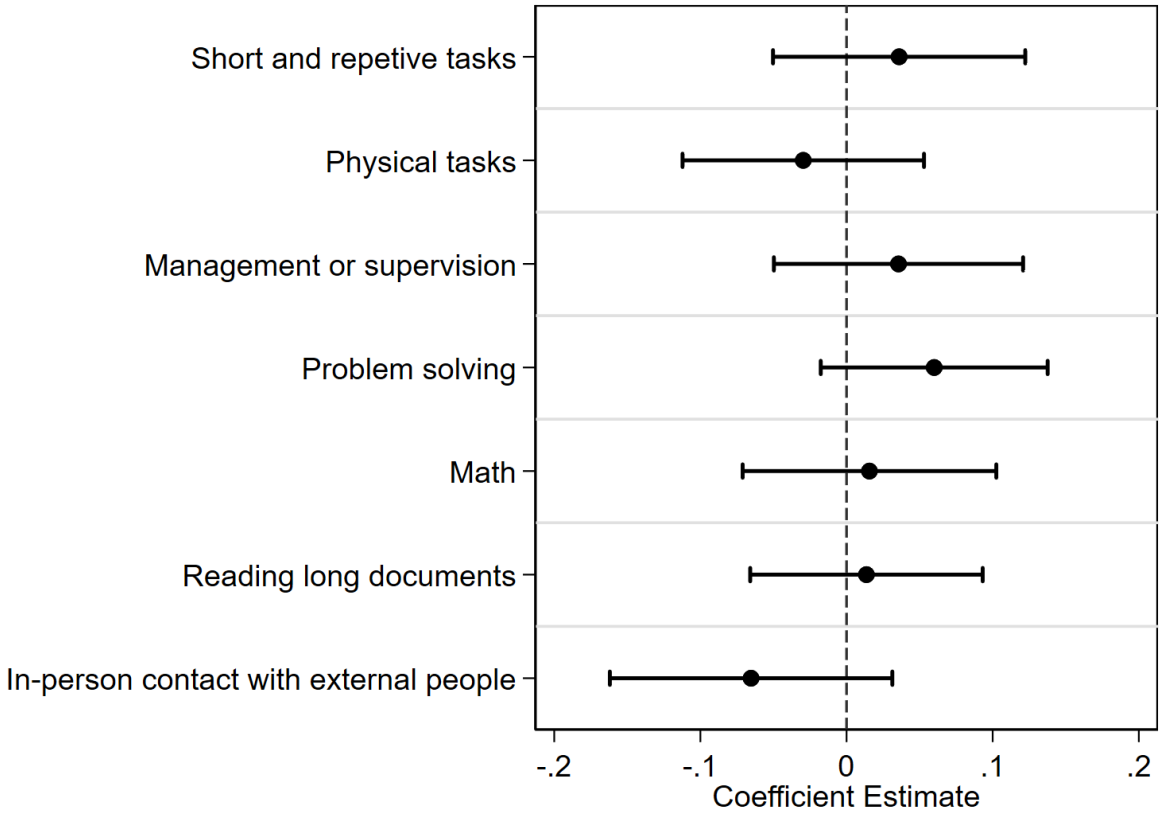
Notes: The figures plot the average share of workers who are NC signers against the return to experience for each occupation and industry. Sample pools the bi-annual 2017–2021 cross sections of the NLSY97. Each circle is proportional to the size of the occupation/industry. A line is fitted and the slope and standard error of the line are weighted by the sizes of occupations/industries.

Figure E3: The Effect of Signing an NC: Robustness to Firm-Level Covariates



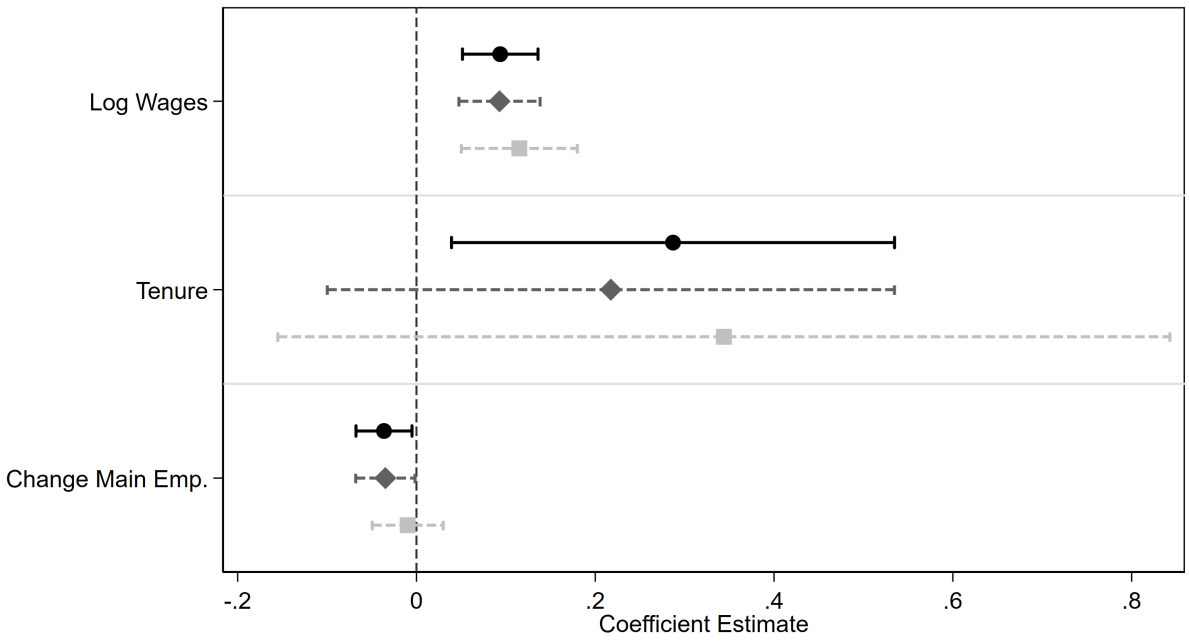
Note: Coefficient estimates are from stacked difference-in-differences estimation, aggregated over post-treatment years, over a bi-annual sample period of 2013-2021 and using cohorts $c \in \{2015, 2017, 2019, 2021\}$. The treatment group for cohort c are those who we observe first signing an NC in year c. The control group consists of workers whom we do not observe holding an NC during the event window and who also changed jobs between year c and the preceding survey year. Job mobility is defined as changing main employers between the current and preceding survey year. The black circle markers report baseline estimates from equation (9). The gray diamond markers with dashed lines report estimates controlling for industry fixed effects. The square markers further add occupation fixed effects. Standard errors are clustered by worker and confidence intervals are reported at the 95% level.

Figure E4: Balance Tests for Changes in Task Content



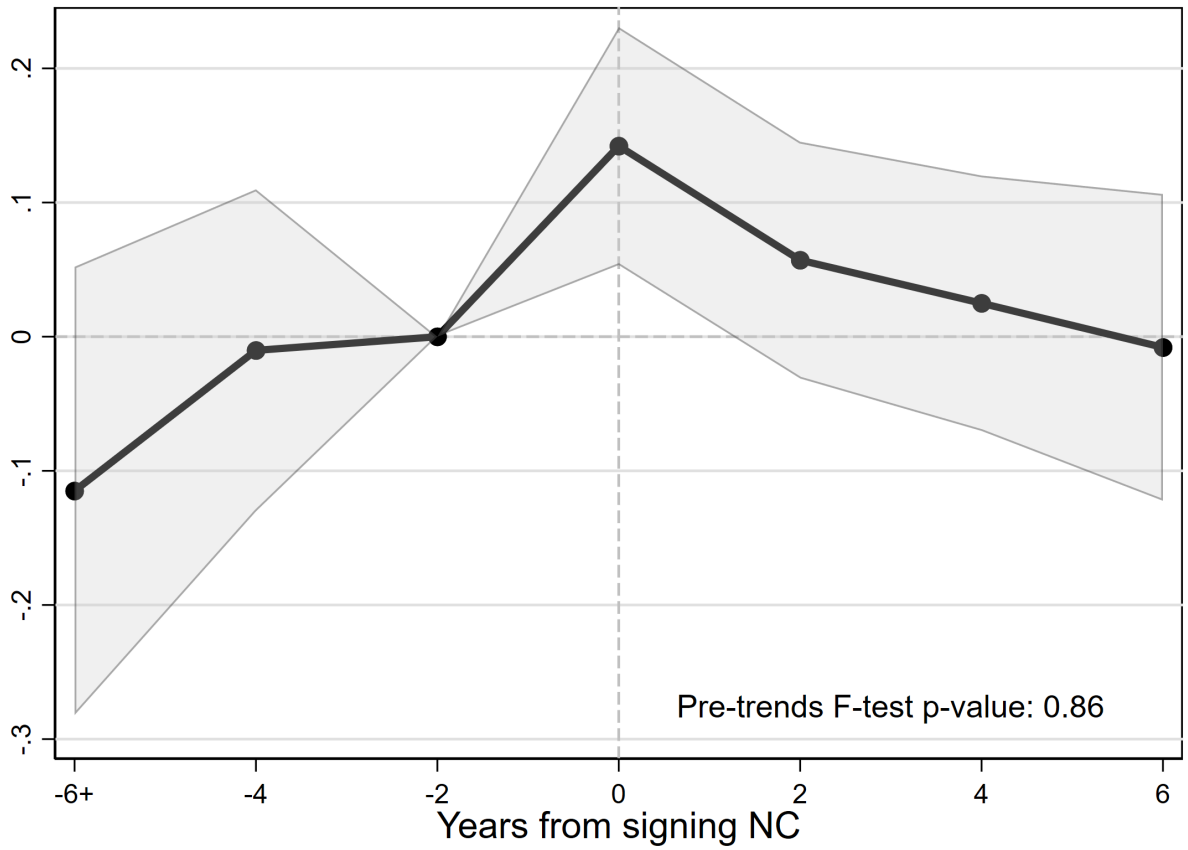
Note: Coefficients report difference-in-difference estimates on task content following equation (9). The outcome variables are indicators for whether a certain type of task is a frequent part of a workers job. Tasks content questions are available in the 2017, 2019, and 2021 survey years, and therefore we use the 2019 and 2021 cohorts and years 2017-2021 for this analysis. Standard errors are clustered by worker and confidence intervals are reported at the 95% level.

Figure E5: The Effect of Signing an NC: Robustness to Different Samples



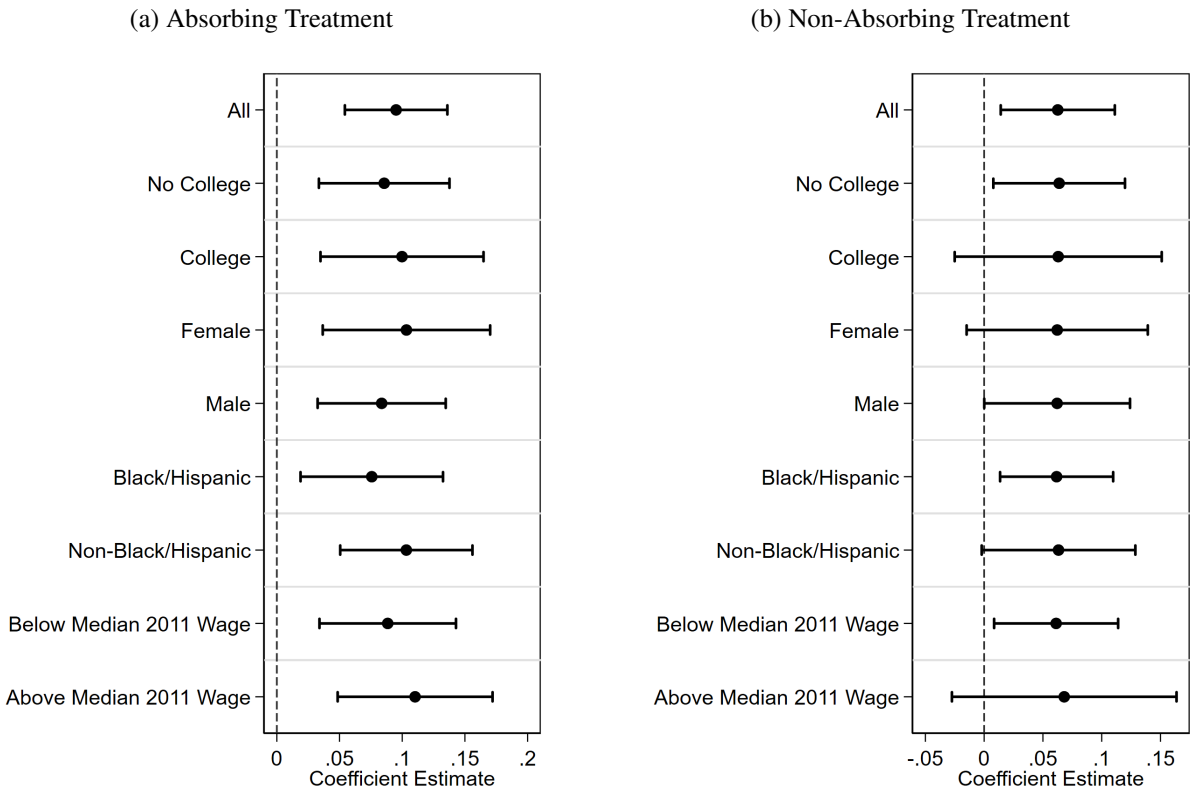
Note: Coefficient estimates are from stacked difference-in-differences estimation, aggregated over post-treatment years, over a bi-annual sample period of 2013-2021 and using cohorts $c \in \{2015, 2017, 2019, 2021\}$. The treatment group for cohort c are those who we observe first signing an NC in year c. The control group consists of workers whom we do not observe holding an NC during the event window and who also changed jobs between year c and the preceding survey year. Job mobility is defined as changing main employers between the current and preceding survey year. The black circle markers report baseline estimates from the main text. The gray diamond markers with dashed lines report coefficient estimates when we restrict attention to years 2015-2021 and cohorts $\{2017, 2019, 2021\}$. The square markers further restrict attention to years 2017-2021 and cohorts $\{2019, 2021\}$. Standard errors are clustered by worker and confidence intervals are reported at the 95% level.

Figure E6: The Effect of Signing an NC on Wages: Later-Treated as Control Group



Note: Coefficient estimates are from stacked difference-in-differences estimation, aggregated over post-treatment years, over a bi-annual sample period of 2013-2021 and using cohorts $c \in \{2015, 2017, 2019, 2021\}$. The treatment group for cohort c are those who we observe first signing an NC in year c. The control group consists of workers who (a) changed jobs between year c and the preceding survey year, (b) do not hold an NC in year c, and (c) sign an NC at some $t > c$ (the later-treated job movers). Standard errors are clustered by worker and confidence intervals are reported at the 95% level.

Figure E7: Wage Effects of Signing an NC from Two-Way Fixed Effects (TWFE) Models



Note: Coefficient estimates report the impact of NCs on wages from two-way fixed effect (TWFE) models. Panel (a) reports coefficient estimates from equation (10) over a bi-annual sample period of 2013-2021. Panel (b) reports coefficient estimates from equation (11) over a bi-annual sample period of 2017-2021. Standard errors are clustered by worker and confidence intervals are reported at the 95% level.