THE LABOR MARKET EFFECTS OF CREDIT MARKET INFORMATION

MARIEKE BOS[§], EMILY BREZA[†], AND ANDRES LIBERMAN[‡]

Abstract. One function of public credit registries is to impose costs on defaulters. This paper exploits detailed data matching credit and labor market outcomes in Sweden and a policy change that provides quasi-experimental variation in the time information on past defaults remains publicly available to document an economically large cost of default in the labor market. When information on past defaults is removed earlier, an individual is more likely to have a job, is less likely to be self-employed, and earns a higher income. The employment cost of default may increase borrower repayment incentives and help sustain uncollateralized consumer credit markets, but may also amplify negative shocks, particularly for vulnerable households.

Keywords: household finance, costs of default, credit information

JEL Classification Codes: G21, G23, D12, D14, J20

1. Introduction

> "Credit reports touch every part of our lives. They affect whether we can obtain a credit card, take out a college loan, rent an apartment, or buy a car — and sometimes even whether we can get jobs ..." – New York Attorney General, 2015.

Credit registries are an important tool used by lenders worldwide to obtain better information about their borrowers and to strengthen repayment incentives. As a result, credit registries are thought to improve the allocation and extent of consumer

Date: March 2016.

We thank Nathan Hendren, Andrew Hertzberg, Emi Nakamura, Matthew Notowidigdo, Wei Jiang, Daniel Paravisini, Thomas Philippon, Nicolas Serrano-Velarde, Daniel Wolfenzon, and Jonathan Zinman, as well as seminar audiences at the Adam Smith workshop, Berkeley Haas, BYU Marriot, Columbia Business School, the Federal Reserve Bank of Philadelphia, MIT Sloan, NYU Stern, Princeton, Bank of Spain, Stockholm University, Riksbanken, and UCLA Anderson for helpful comments. Jesper Böjeryd provided excellent research assistance. Funding from VINNOVA is gratefully acknowledged. All errors are our own.

[§]Visiting Scholar Federal Reserve Bank of Philadelphia, Stockholm University. E-mail: `marieke.bos@sofi.su.se`.

[†]Columbia University. E-mail: `ebreza@gsb.columbia.edu`.

[‡]New York University. E-mail: `aliberma@stern.nyu.edu`.

credit (Djankov, McLiesh, and Shleifer (2007)). Multilateral institutions such as the International Monetary Fund and the World Bank urge countries to adopt registries, citing them as a fundamental step toward financial development. Indeed, several studies have documented that this information affects borrowers' access to credit.[1]

However, much less is known about the effects of credit market information on noncredit outcomes, such as employment, that are critical for welfare and policy analysis. Credit information may impose an employment cost of default indirectly through its effects on credit supply, but more direct channels are also possible. Namely, if nonbank actors such as employers make decisions based on credit information, then a signal of a past default itself may affect noncredit outcomes. In particular, insurance companies, utilities, landlords, mobile phone providers, and other service providers typically check an individual's credit history before entering into long-term contracts with them. There is also ample anecdotal evidence that many employers around the world query credit registries when making hiring decisions.[2]

We begin by documenting a large raw correlation between credit scores and future employment using a representative sample of the entire adult population of Sweden between 2000 and 2005.[3] An individual whose estimated 12-month probability of default is 10% higher has a 0.3% higher probability of being unemployed the following year. The goal of this paper is to understand whether part of this relationship is causal.

We use detailed credit registry, labor market, and tax data from Statistics Sweden alongside a natural experiment that changed the amount of time that records of past delinquencies were retained on consumer credit reports to show that bad credit information following default negatively affects an individual's future employment. To our knowledge, we are the first to highlight this interlinkage between the labor market and credit information and to measure the employment costs of default.

[1]For example, see Musto (2004), Brown and Zehnder (2007), De Janvry, McIntosh, and Sadoulet (2010), Bos and Nakamura (2014), González-Uribe and Osorio (2014), Liberman (2015).

[2]In the U.S. 47% of the firms check the credit information of their prospective employees according to: http://www.shrm.org/research/surveyfindings/articles/pages/creditbackgroundchecks.aspx. According to estimates obtained from the leading credit bureau, in Sweden, roughly 15% percent of all inquiries to the credit registry are made by nonfinancial institutions conducting background checks of potential employees. These non-financial institutions employ approximately 37% percent of the Swedish labor force. The Swedish Government Employment Agency lists jobs that currently require a clean credit record: financial, transportation, real estate, retail, and security (See http://www.arbetsformedlingen.se).

[3]Note that in Sweden everyone, including individuals without any credit history, is registered at the credit bureau from age 17 onward.

When a borrower defaults in Sweden, she receives both an arrear and a nonpayment flag that appears at the top of her credit report.[4] Swedish law establishes that each arrear must be deleted three years after the delinquency occurred. The nonpayment flag at the top of the report remains until all of the arrears have expired. Before October 2003, arrears were deleted on the last calendar day (i.e., December 31) of the third year after the nonpayment. Beginning in October 2003, the law was reinterpreted, and arrears were deleted exactly three years to the day after they were generated. Importantly for identification, the key impetus for this change was technological and coincided with the upgrade of the computer systems used by the registry. We note that this policy change was first exploited by Bos and Nakamura (2014), who find that shorter retention times result in an increased supply of credit. However, our research question and identification strategy differ significantly from Bos and Nakamura (2014).

A schematic representation of the policy change is shown in Figure 1. Consider, for example, an individual who defaulted in February 2000. Note this individual was minimally affected by the policy change, and therefore the record of her default was publicly available in the credit bureau until the end of September 2003, three years and eight months later. Next, consider an individual who defaulted in February 2001. She was more affected by the policy change. The record of her default was publicly available in the credit bureau only until February 2004, exactly three years later. Thus, defaulting in February of 2000 or February of 2001 led to different retention times of the nonpayment flag, namely an eight-month reduction in the retention time of past defaults on the credit bureaus for the February 2001 defaulter relative to the February 2000 defaulter. Given that the policy change was announced in March 2003, all individuals who defaulted in 2000 or 2001 did so under the same beliefs about arrear retention time.

We use the variation in the retention time of the nonpayment flag caused by the policy change to identify the causal effect of credit information on employment outcomes. For our analysis, we use the population of Swedish individuals who ever borrowed from a pawn-broker. This sample of alternative credit borrowers is well suited for our analysis, because these individuals are more likely to face periods of

[4]The credit bureau in Sweden also provides a continuous credit score, which is affected by nonpayment flags. We focus on the binary nonpayment indicator in our main analysis. This nonpayment flag is what employers can observe from an individual's credit record and has a discrete adverse impact on an individual's credit score. This is similar to the US, where employers can view trade lines and information about specific non-payments, but are not allowed to observe the FICO score.

financial distress that lead to the reporting of nonpayments on their credit reports. Furthermore, exclusion from the the labor market is likely to be quite costly for this population.[5] We refer to the 2001 cohort of defaulters in our sample as the Treated group and the 2000 cohort of defaulters as the Control group. The policy caused a decrease in the average retention time of past defaults for members of the Treated group relative to the Control group. A simple comparison of Treated and Control individuals before and after the removal of their respective past default information would confound any causal treatment effect with other annual trends in the Swedish economy. Instead, we take advantage of the fact that the policy change affected the retention time of the past default indicator in a differential manner for individuals who defaulted in different calendar months. In our example above, a Treated individual who defaulted in February experienced an eight month reduction in the retention time of her past default relative to a Control individual who also defaulted in February. However, because the policy went into effect in October, 2003, Treated and Control individuals who defaulted in October, November, or December experienced the same retention time of exactly three years. Thus, in our main empirical strategy we compare employment outcomes for individuals in the Treated and Control groups who received a nonpayment flag early in the year (February to May) with those who received an arrear late in the year (August to November). We track how these outcomes change after the nonpayment flag is deleted (i.e., four and more years after default) relative to the first three years after default.[6] The coefficient of interest can then be interpreted as the causal effect of the removal of information on past defaults on employment and other labor market outcomes.

We find that the removal of information on past defaults has large employment effects. A Treated individual who defaulted early in the calendar year is approximately three percentage points more likely to be employed the year in which their nonpayment information is removed from the credit registry, relative to a Control individual,

[5]While in principle we could use the representative sample of the entire Swedish population to conduct our analysis, there are at least two reasons why in practice we chose not to do so. First, the size of the representative sample after we impose all restrictions, in particular, to individuals that defaulted in a particular year, is roughly 300 individuals, which seriously compromises the power of our tests to detect any effects. Second, the credit bureau data was delivered without a social security number, which makes it impossible to match to most of the labor market outcomes used in our analysis.

[6]In addition, we restrict our sample to those individuals who did not default again in the subsequent 20 months. This restriction ensures that individuals are not classified simultaneously in multiple treatment groups and improves the power of our tests. Note that both the default and repayment decisions that affect treatment status were made before the announcement of the Swedish policy change.

and relative to an individual who defaulted late in the year. This difference persists (at least) one year after the information is removed from the registry, albeit with a smaller magnitude. Importantly, consistent with our identification assumption, we find a positive monotonic relationship between the size of the reduction in retention time –e.g., eight months for the February defaulters, seven months for the March defaulters, etc.– and the probability of being employed.

Further, we find that individuals whose information is removed earlier earn higher incomes, are less likely to be self-employed, are less likely to pursue additional years of education, and are more likely to change residence. These effects are stronger among those with fewer years of schooling. They are also strongest for individuals who were unemployed before arrear deletion, but who had earned some wage income in the prior three years. There is no detectable effect for the so-called "chronically unemployed" – those individuals who had not earned any labor income in the three years preceding arrear deletion. Individuals employed in the preperiod also experience a positive effect. Finally, we find that the employment effects are strongest in regions (kommuns) with below-median levels of baseline unemployment.

Public information on past defaults may affect employment outcomes through many channels. First, removing the record of past defaults increases an individual's access to credit. This is documented by Bos and Nakamura (2014) in our setting. In turn, this increased access to credit may impact employment by allowing individuals to make investments necessary for finding (i.e., by allowing the individual to search) and keeping (i.e., by making the individual more productive) a job (e.g., (Karlan and Zinman (2009); Mullainathan and Shafir (2013); Kehoe, Midrigan, and Pastorino (2014)). Increased credit may also allow individuals to invest and become entrepreneurs, thus reducing the relative value of wage labor (e.g., Chatterji and Seamans (2012); Hombert, Schoar, Sraer, and Thesmar (2014); Greenstone, Mas, and Nguyen (2014); Schmalz, Sraer, and Thesmar (2015); and Adelino, Schoar, and Severino (2015)). Further, if individuals use labor hours to smooth negative shocks in a precautionary manner, they may reduce their labor supply following an increase in credit supply (e.g., Low (2005); Pijoan-Mas (2006); Jayachandran (2006); and Blundell, Pistaferri, and Saporta-Eksten (2016)). Note that both the entrepreneurship and labor smoothing channels are inconsistent with our baseline effect, by which more access to credit leads to more wage employment.[7] Information may also affect employment more directly if employers screen applicants directly through credit checks or indirectly through other traits that depend on credit information, which may in turn augment any productivity effects from an increase in credit supply.[8]

The estimates of the employment cost of default we present do not depend on the specific channel that may generate them. However, our preferred interpretation of the mechanisms driving the findings is that they combine an increased ability to invest in job search and productivity, mediated through an increase in credit supply, together with an effect through employer screening.[9] We present four pieces of evidence that suggest that the increase in credit supply is not able to explain our results by itself. First, the effect of the removal of information of past defaults on the individual's spouse's income is not significantly different from zero, while the point estimate is even negative. This pattern of intra-household spillovers is less consistent with household-level credit constraints that limit necessary search and productivity investments from taking place. Second, individuals are less likely to be self-employed following removal of their negative credit information, despite their increased access to credit. If the investments required for self-employment activities are of the same order of magnitude as those required for search and productivity investments, this suggests that credit constraints may not fully prevent these investments from being made. Third, we find that the number of credit checks by non-financial institutions remains constant after information on past defaults is removed, suggesting no increase in an individual's search intensity. Fourth, we measure large employment effects in tighter labor markets but undetectable effects in areas with high unemployment.[10] At the same time, we find large effects of credit information on access to credit in both areas. Thus, the relaxation of credit constraints alone is not sufficient to induce an effect in observed equilibrium employment and suggests a role for employer screening through credit information.

[7]Our results are also inconsistent with Herkenhoff (2013) and Herkenhoff and Phillips (2015), who study a matching model of the labor market, where access to credit leads to higher unemployment through an increase in the employee's outside option.

[8]For evidence of employer screening along other dimensions see e.g. Kroft, Lange, and Notowidigdo (2013) (unemployment spell) and Deming, Yuchtman, Abulafi, Goldin, and Katz (2016) (postsecondary school).

[9]Screening by landlords may also contribute to the causal effect of information on employment by affecting mobility. We perform a bounding exercise and show that increased mobility following the removal of credit information can explain at most a quarter of the magnitude of our results.

[10]This result is consistent with Kroft, Lange, and Notowidigdo (2013).

By highlighting and measuring the previously ignored employment cost of default, our results inform the academic and policy debates about the appropriate level of the costs of default and the design of bankruptcy policies to sustain repayment in consumer credit markets (Chatterjee, Corbae, Nakajima, and Ríos-Rull (2007); Livshits, MacGee, and Tertilt (2007)). The employment cost of default may improve ex-ante repayment incentives, potentially leading to a more efficient allocation of credit. Moreover, employers may experience productivity gains by using information about debt repayment to screen workers and improve matching in labor markets (e.g., Autor and Scarborough (2008)). However, ex post, these costs of default may be borne disproportionally by poorer households that are more exposed to negative shocks. Credit information may thus induce a multiplier effect on unemployment and a negative spiral from which households may be unable to break free.[11] Further, higher costs of default may reduce a household's demand for credit and thus its ability to smooth consumption in bad states. The welfare implications of the credit-labor market linkage are thus ambiguous and beyond of the scope of this paper.

Our results also speak to the current policy debate surrounding the appropriate scope for use of credit information by employers. In April 2015, the New York City Council voted to dramatically restrict credit checks in hiring, and similar bills have been passed in 10 states and in Chicago.[12] Indeed, data and information technologies, i.e. "big data", are likely to become a prevalent feature of modern economies, while the tradeoffs involved have received scant attention in the academic literature (e.g., see Einav and Levin (2013)). Our results provide timely inputs to begin to evaluate these tradeoffs.

Our work contributes to several strands of the household finance literature. First, we contribute to the empirical and theory literatures on the impacts of credit market information on credit market outcomes.[13] Second, we add to the literature that studies the effects of debt renegotiation on households (e.g., Dobbie and Song (2015);

[11]In aggregate and under some conditions, our findings suggest an avenue complementary to Mian and Sufi (2010) through which large debt build-ups followed by financial distress may result in large fluctuations in consumption.

[12]The New York City bill does contain multiple exemptions, including for example for police officers, employees with state or federal security clearance, and workers with access to third-party assets of more than $10,000. See New York Attorney General (2015), and press reports such as http://www.nytimes.com/2015/04/17/nyregion/new-york-city-council-votes-to-restrict-credit-checks-in-hiring.html.

[13]The empirical evidence includes Musto (2004), De Janvry, McIntosh, and Sadoulet (2010), Bos and Nakamura (2014), and González-Uribe and Osorio (2014). Theoretical papers include Pagano and Jappelli (1993), Padilla and Pagano (2000), and Elul and Gottardi (2015), among others.

Liberman (2015)). Third, we contribute to the literature on the interaction between entrepreneurship and credit supply (e.g., Chatterji and Seamans (2012); Hombert, Schoar, Sraer, and Thesmar (2014); Adelino, Schoar, and Severino (2015); and Banerjee, Breza, Duflo, and Kinnan (2015)). Finally, we also contribute to a literature that studies amplification of negative shocks (e.g., Agarwal, Chomsisengphet, Mahoney, and Stroebel (2015); Kroft, Lange, Notowidigdo, and Katz (2015)).

The remainder of the paper is organized as follows. Section 2 describes the data, setting, and empirical strategy. In Section 3 we present the results. In Section 4 we show additional tests that suggest that employer screening is likely to explain part of our results. Section 5 concludes.

2. Measuring the employment cost of default

We begin by documenting a strong observational relationship between loan repayment behavior and employment status using a panel data set of a random sample of 15,862 individuals from the Swedish population, matched to tax records. Table 1 presents the output of an OLS regression of $1\,(wages_{i,t} > 0)$, a dummy that equals one for individuals (indexed by i) with any positive wage income during a year (indexed by t), on one- and two-year lagged logarithm of credit scores, $log\,(creditscore_{i,t-l})$ ($l = 1$, $l = 2$), with controls X_i, which include demographic characteristics (columns 1 and 2) or individual fixed effects (columns 3 and 4):

$$1\,(wages_{i,t} > 0) = \alpha + \beta_l log\,(creditscore_{i,t-l}) + \gamma X_i + \varepsilon_{i,t}.$$

In Sweden, higher credit scores are indicative of worse repayment behavior (the score reflects the individual's default risk on a scale from 0 to 100). Hence, the negative coefficients on lagged credit scores in columns 1 and 2 indicate that worse credit histories are associated with a lower probability of being employed in the future. Further, as columns 3 and 4 show, this result is true even within an individual's own history.

This strong correlation is likely driven by many effects. First, individuals who lose their jobs and remain unemployed may causally have a higher propensity to default on their debts (reverse causality).[14] Second, individuals who are more likely to be unemployed may also be the types of people who are more likely to default on their debts and have a signal of nonpayment in their records (omitted variables). In this paper we study whether at least part of this correlation, additionally, reflects a causal

[14] For example, see Foote, Gerardi, and Willen (2008a) and Gerardi, Herkenhoff, Ohanian, and Willen (2013).

effect of negative credit information on employment. To do this, we exploit a plausibly exogenous source of variation in the credit information observable through the credit bureau, holding the two other effects constant.

In what follows we turn to our empirical tests to uncover the direction and magnitude of the causal relationship of credit information on labor outcomes.

2.1. **Setting and policy change.**

Swedish credit bureaus and policy change. Credit bureaus are repositories of information on the past repayment of debts and other claims, such as utility bills, credit cards, and mortgage payments (Miller (2000)). In Sweden, credit bureaus collect registered data from three main sources: the national enforcement agency (Kronofogden), the tax authorities, and the Swedish banking sector. Banks typically report a borrower to be in default when 90 days past due. Other entities with access to the credit bureau, such as phone companies, exercise discretion when a consumer is reported as delinquent.[15] Each reported default triggers an arrear on the borrower's credit report. Furthermore, in Sweden any individual or company can view the credit records of any individual.[16]

Before October 2003, Swedish law mandated all arrears to be removed from each individual's credit report three years after the nonpayment occurred. In practice, the credit bureau removed all arrears on December 31 of the third year after the nonpayment occurred. In 2003, the Swedish government changed the interpretation of the law so that every arrear would be removed from the credit bureaus, and thus no longer be publicly available, *exactly* three years after the day the nonpayment was recorded. This change was motivated by an upgrade to the bureau's IT capabilities and a reduction in the cost of distributing information. This law was implemented in October of 2003.[17]

As shown in Figure 2, the adjustment to the law induced a sharp change in the pattern of removal of arrears by the credit bureaus. The figure plots the number of individuals whose arrears are no longer reported in the credit bureau at a bimonthly frequency. The figure shows that before 2003, arrears were only removed from the

[15]Individuals have the option of filing a protest to the courts to correct potential errors.

[16]The law states that an individual's credit records are available to other parties as long as the parties plan to enter into a contractual relationship with the individual.

[17]The Swedish government announced their decision to change Paragraph 8 of the law that regulates the handling of credit information (KreditUpplysningsLagen or credit inquiry law) on July 2003, and the law change went into effect on October 2003. See http://rkrattsdb.gov.se/SFSdoc/03/030504.PDF

credit registry on the last day of the year. In our bimonthly data, an individual who had an arrear on December 1st, but had that arrear removed on December 31, is first observed without an arrear in February. Thus, the figure shows that before 2003, the vast majority of arrears were removed only once per year in December (i.e., corresponding to the February spikes in in our data). Further, the figure shows a noticeable spike in the frequency of removals in October 2003. This spike corresponds to the removal of the stock of arrears that had occurred between January and the end of September 2000 and that had not yet been deleted from the credit bureau. After October 2003, the frequency is more smoothly distributed over the year, in effect following the distribution of nonpayments during the year, three years earlier.

Identification intuition. We attempt to identify the causal effects of variation in past repayment information on employment and other labor market outcomes. An idealized experiment to identify this effect would consider two identical groups of individuals –Treated and Control– who have defaulted in the past and as a result have a bad credit record. In that experiment, the credit bureau would delete the information for the Treated group earlier than scheduled, and any subsequent difference in employment between the two groups would be causally assigned to this change.

In our empirical setting, we use the variation in the retention time of publicly observable arrears induced by the 2003 policy change in Sweden to approximate this idealized setting. One naive empirical strategy would be to focus on cohorts before the policy change and to compare individuals who defaulted earlier in the year to those who defaulted later in the year. After all, the early defaulters did experience longer retention times than the end of year defaulters. However, it is likely that individuals who default at different times during the year differ in ways that may be correlated with labor market outcomes. Further, individuals may have been aware of this feature of credit bureaus and chose to time their defaults accordingly. Hence, a comparison of the employment prospects of individuals who defaulted early and late in the same year before the policy change is likely to be biased.[18]

Instead, the policy change induced *unexpected* variation in the length of time that information was retained in the credit bureaus. Hence, individuals who defaulted

[18]In Online Appendix Table IAI we show the mean of a set of variables for individuals whose last default occurred early (in February, March, April, or May) and late (in August, September, October, or November) in our sample, and note that for all but one of these variables, the difference between the means of the two groups is significant at the 1% level (the only variable in which the means are not significantly different is "financial inquiries", which measures the number of inquiries to the individual's credit records by financial institutions).

in 2000, three years prior to the policy change, did so under the same beliefs about retention time as individuals who defaulted in 2001, two years before the policy change. The unexpected nature of the policy change allows us to rule out any strategic behavior of individuals timing their default so as to experience shorter retention times.

An alternative identification strategy is to compare individuals who defaulted in 2000, which we define as the Control group, with those who defaulted in 2001, which we define as the Treated group, observing that the average retention time is lower for the Treated group. However, this strategy is also problematic as there may be other differences between individuals who defaulted in 2000 and 2001 that may be correlated with labor market outcomes.[19]

Instead, we combine the two empirical strategies –Treated versus Control cohorts and early versus late defaulters within the calendar year– for identification. We compare the difference in the employment prospects of individuals whose default was reported early and late in the year 2001 (Treated), with the same difference but for individuals whose default was reported the previous year, 2000 (Control). We observe that individuals in the Treated group (who defaulted at any point in 2001) or individuals in the Control group (who defaulted late in 2000) were subject to the same three-year retention times. Individuals in the Control group who defaulted early in the year, say in March, were subject to three years plus seven months of retention time. This double-difference analysis is the basis of our identification strategy. We then take a third difference and compare outcomes for each individual before and after the three-year post-arrear date. The identification assumption we make is that, in the absence of the policy change, the difference in employment outcomes of individuals in the Control and Treated groups whose defaults were reported early and late in the year would have remained constant before and after the deletion of the nonpayment flag. In Section 3.3.1 we provide evidence that is consistent with this assumption.

Finally, note that within the group of Treated individuals, those who defaulted earlier in the year experienced a larger decrease in retention time than those who defaulted later in the year. This suggests an additional test of our identification

[19]Another thought might be to try to identify such effects using the fact that in the US, Chapter 7 bankruptcy information is removed from the credit report after 10 years. Indeed, Musto (2004) finds a large jump in FICO score and in credit supply following bankruptcy flag deletion. However, it is important to note that US employers are not allowed to observe FICO scores. Further, deletion from the credit report does not coincide with deletion from the public record. Thus, if past bankruptcy information appears on other types of materials collected during employee background checks, this deletion may not even have an impact on the information available to the employer.

strategy: the effects of credit market information on employment should be monotonically decreasing in the time of the year during which individuals' defaults were initially reported. In Section 3, we provide evidence that is consistent with this intuition.

Next, we describe our data and detail how we implement our empirical strategy.

2.2. **Data.** Our sample corresponds to the universe of borrowers of alternative credit in Sweden. This sample was generously supplied by the Swedish pawnbroker industry and contains information about the 132,358 individuals who took out a pawn loan at least once between 1986 and 2012. This sample is well suited for our analysis. Indeed, these individuals are more likely to face periods of financial distress that lead to the reporting of nonpayments on their credit reports compared with the general population. Furthermore, this group of individuals represents a group with lower levels of income and education than the general population, and exclusion from the the labor market is likely to be quite costly.[20]

Our credit data correspond to a panel at a bimonthly frequency, with observations from 2000 to 2005. We observe a snapshot of each individual's full credit report from the leading Swedish credit bureau, Upplysningscentralen. Unlike in the U.S., Swedish credit bureaus have access to data from the Swedish Tax authority and other government agencies. This enables us to observe, in addition to all their outstanding consumer credit and repayment outcomes, variables such as home ownership, age, marital status, yearly after- and before-tax income from work, and self-employment. Importantly for this study, we observe when an individual's nonpayment was first reported and when it was removed by the credit bureau.

To obtain labor market outcomes, we match the credit bureau data with information obtained from Statistics Sweden (SCB). These data are at the yearly level, for the years 2000-2005 and include information on each individual's employment status. This status can take one of three categories: employed, defined as fully employed during the entire year, partially employed, defined as having been previously unemployed during the year, and not employed. The data also include measures of income such as pretax income, wages, and income from self-employment. We defer an analysis of summary statistics of our main outcome variables until after we have presented our sample selection criteria.

[20]See Bos, Carter, and Skiba (2012) for a comparison of the sample to both the Swedish and US populations.

2.3. **Implementation of empirical strategy.** To isolate those individuals who were most directly affected by the policy change, we make three sample restrictions. First, we include in our analysis sample only individuals who received an arrear for nonpayment in 2000 or 2001 and thus had those nonpayment flags removed in 2003 or 2004. Second, we further restrict the sample to those individuals who did not receive additional arrears in the subsequent two years (all before the policy change). Note that all individuals in our final analysis sample made their nonpayment (and subsequent payment) decisions under the same beliefs about the Swedish credit registry data retention policies. Thus, the actions that caused an individual to fall into our analysis sample are predetermined relative to the policy change. Our a priori hypothesis is that individuals will have the greatest change in outcomes when their last arrear is erased from the information registry. Thus, this second sample restriction criterion allows us to approximate this group of individuals using predetermined decisions. Third, because of the bimonthly nature of the credit registry data shared with the researchers (e.g., December-January defaulters are first reported in the February snapshot, February-March in the April snapshot, etc.), we restrict our sample to defaults occurring strictly after January 2000.[21] The December-January bimonthly flow of removals considers individuals whose information was deleted as of the previous year (close to exactly three years after it occurred), and as such will distort our estimates. For the same reason, we omit individuals whose defaults are removed from the credit bureau in the December-January 2001 bimonth. Finally, we focus on individuals who are between 18 and 75 years old the year before information on past defaults is removed from the credit registry. These selection criteria result in a sample of 15,232 individuals.

Figure 1 depicts the timeline of the policy change and how it affected the length of time in which nonpayments were reported for the individuals in our sample. In particular, Control group individuals whose nonpayment was recorded in the first months of the year were reported in the credit registries for a maximum of almost three years and eight months until the end of September 2003, while Treated group individuals whose nonpayments were recorded in the first months of the year were reported in the credit registries for exactly three years. Figure 1 also shows the number of past defaulters in each of the bi-monthly bins. We note that while in

[21] Note that the credit bureau updates its information on a daily basis. The research team, however, was only allowed access to bimonthly snapshots of the data.

both cohorts there are substantially more early defaulters than late defaulters, these patterns are remarkably consistent between the Treated and Control groups.

We define the Treated groups with the variable $treated_i$, which equals one if borrower i's last nonpayment occurred during 2001 and zero if it occurred during 2000. We interact $treated_i$ with the dummy variable $early_i$, which distinguishes between individuals whose nonpayments occurred early and late during the year. Because in our data each individual is assigned to a bimonthly cohort of defaulters, we define $early_i$ to equal one for those individuals whose last nonpayment occurred in February-March or April-May, and zero for individuals whose last nonpayment occurred in August-September or October-November.[22] Finally, we create a dummy called $post_{i,t}$ which equals one for all event years after borrower i's nonpayment signal is removed (2003 for the Control group, 2004 for the Treated group). Note that the variable $post_{i,t}$ is measured in event time, which is normalized to zero in 2000 for the Control group and in 2001 for the Treated group. Thus, event time year three represents the year in which the nonpayment flag is deleted from the credit bureau for any individual in our sample. We include individual fixed effects ω_i, year fixed effects ω_t, and event time fixed effects ω_τ, as well as all double interactions that are not absorbed by fixed effects. Our main specification is the following reduced form model:

$$\begin{aligned} employed_{i,t} \;=\; & \omega_i + \omega_t + \omega_\tau + \beta treated_i \times early_i \times post_{i,t} + \delta post_{i,t} + \\ & \gamma treated_i \times post_{i,t} + \lambda early_i \times post_{i,t} + \varepsilon_{i,t}. \end{aligned} \tag{2.1}$$

Note that ω_i absorbs the baseline and interaction coefficients of $treated_i$ and $early_i$. The coefficient β, which is our main outcome and which we report with our regression output, measures the differential probability of being employed for the Treated and Control group, for individuals whose nonpayment was reported early in the year relative to those whose nonpayment was reported late in the year, the year(s) after each individual's nonpayment is no longer reported relative to the three prior years. The coefficients δ and λ capture differences in employment for individuals whose nonpayment occurred late- and early-in-the-year, respectively, the years after the arrear is no longer publicly available. Finally, γ captures differential employment trends for all Control group individuals after their nonpayment information is no longer publicly available.

[22]Note that to make the early and late groups comparable in size we exclude the June-July cohort. However, below we do include individuals in this cohort when we measure differential effects by differential intensity of the treatment by month of nonpayment.

We note that no single sufficient statistic captures all possible effects of credit information on labor market outcomes. A priori, one candidate might be the credit score. However, non-financial actors do not observe credit scores and instead only observe non-payment flags or arrears. Further, even banks may incorporate non-payment flags above and beyond their contribution to the credit score in their internal credit models. In the Online Apendix Section B we present the results of an exercise where we use our triple interaction $treated_i \times early_i \times post_{i,t}$ as an instrument for credit score. However, we prefer and show throughout the paper the reduced form ITT specification (2.1).

2.4. **Summary statistics.** Before presenting the regression output, we present in Table 3 selected summary statistics of our outcome variables. We focus our analysis on employment outcomes, broadly construed. In addition to earnings and whether an individual has a job, we also consider alternatives to labor income, including seeking more education and turning to self-employment income. The top panel presents a brief definition for each of our outcome variables. In turn, the lower panel displays selected summary statistics of these outcome variables for our estimation sample. Our sample includes 15,232 individuals, which we observe for six years (2000 to 2005). Our summary stats are estimated the three years before these individuals' nonpayment flags are removed, which correspond to 2000 to 2002 for the Control group and 2001 to 2003 for the Treated group. During those years, an average of 43 percent of individuals in our sample are employed during the full year, while 79 percent received some positive wage income. We transform our income measures, which are in units of hundreds of Swedish Kronor (SEK), to logarithms and add a 1 to include the effect of zero income. On average, log(income + 1), the log of pretax income, equals 5.6, which corresponds roughly to a pre-tax income of 102,000 SEK or \$12,200. This figure is roughly 50 percent lower than the population mean during that period. This difference stems mostly from high levels of unemployment in our sample. Further, log(wage + 1), the log of wages, also equals 5.6 on average. Indeed, wages represent almost all of the income these individuals receive. Roughly 5 percent of all individuals in our sample appear as self-employed. Further, 7 percent of individuals change addresses each year (defined as changing kommun, a geographic division akin to an MSA–there are roughly 200 kommuns in Sweden). Finally, individuals in our sample are 42.8 years old, 60 percent male, and nine percent own a house.[23]

[23]In the 2001-2003 period, these numbers were 40.6 years old, 49.5 percent male, and 59.9 percent home ownership, respectively, for the entire population of Sweden.

3. The Effect of Credit Information on Employment

In this section we present and discuss our main results. We start by showing graphically the event-time evolution of the average outcomes, which provides evidence in support of our identification assumption.

3.1. **Graphical evidence.** The identification assumption for regression (2.1) is that, in the absence of the regime shift, the probability of being employed for the Treated and Control groups, between early- and late-in-the-year defaulters would have evolved in parallel. We provide evidence that supports this assumption in Figure 3. The top panel shows the average of $employed_{i,t}$, which is a dummy that equals one if the individual was employed during that period, as well as the average of a dummy that equals one for individuals who receive any positive wage during the year. The x-axis shows event time years, which are defined starting at zero in 2000 for the Control group and in 2001 for the Treated group. There are no detectable differences in the trends of the difference of either variable between early and late defaulters in the Treated and Control groups during the three years before removal of the nonpayment flag (i.e., in event times 0 to 2). Similar effects can be observed for the average log income and log wage income, where zeros have been replaced by ones, shown in the lower panel. These graphs provide evidence that is consistent with our identification assumption. The figures also hint at our main results: Treated group individuals who defaulted early in the year exhibit a higher probability of employment and earn higher incomes after their nonpayment flags are removed.[24]

3.2. **Main results.** Table 3 presents the coefficient of interest of specification (2.1). In columns 1 through 4, we study the effect of shorter retention times on employment defined in two different ways. First, columns 1, 2, and 3 present the regression results when the outcome is *employed,* defined as a dummy for whether the individual was employed either part time or full time throughout the year. Column 1 documents that the probability that an individual whose information is reported for a shorter

[24] One potential concern with our identification strategy is that individuals in the Treated group who defaulted early in the year are somehow different to Control individuals who defaulted early (relative to late defaulters) in a way that only manifests itself in event time 3 and later, for reasons that are unrelated to their credit information. For example, if these groups of individuals exhibit different propensities to be employed according to the business cycle, and the business cycle evolved in a way that produced exactly this time series pattern. In Online Appendix Table IAII we present the evolution of a set of Swedish macroeconomic variables between 1999 and 2005. There appear to be no such obvious relationship in the macro data: for example, during this period there are no recessions, and unemployment only varies between 5.2% and 7.7%.

period is employed increases by 2.8 percentage points the year their nonpayment is removed from the credit registry (year three). This effect represents a 6.5 percent increase relative to the preperiod average employment rate (43 percent). Column 2 shows that this effect is also significant for the two years after removal on average, although with a lower magnitude. Column 3 shows that focusing only on the second year after removal, the point estimate continues to be positive, although statistical significance is lost.

Columns 4, 5, and 6 of Table 3 show that the same pattern emerges when employment is defined instead as receiving any positive labor market income during the year. Indeed, Column 4 shows that Treated group individuals who defaulted early in the year are 3 percentage points more likely to earn positive income from work, and this effect persists two years after the information was removed. Furthermore, the probability of receiving positive income from work is positive (and statistically significantly so) and of the same magnitude during the second year, as shown in column 6.[25]

We further exploit our empirical setup to explore the impact of credit market information on additional labor market outcomes. Columns 1 to 3 of Table 4 display the output of our main regression model (2.1) during the two years after the removal of the nonpayment flag for an array of additional labor market outcomes, including the log of income from work, log(wage + 1), the probability of being self-employed, and the log of total pretax income, log(income + 1). Income measures are in hundreds of SEK. To capture the effects of zero income from employment in the logarithms, we replace zeros with a one.[26]

Consistent with our previous results, we find that individuals whose nonpayment flag was retained for less time earn higher wage incomes. As per the summary stats

[25]To get a sense of the economic magnitude of these effects is by linking the direct effect of credit information on credit scores with the effect on employment. Note that average credit score is lower for early defaulters in the Treated group in the event year in which their information is deleted (event year 3) because their signal is removed earlier in the year (on average, 6 month earlier) than late defaulters and than Control group defaulters. In the Internet Appendix Table IAV we do so by instrumenting the logarithm of yearly average credit score with our triple interaction variable in regression (2.1). The results of this test, which we interpret with caution given that employers can only observe the flag of past defaults and not the actual credit score, imply that a ten percentage point decline in the yearly average credit score, which in Sweden corresponds to a reduction in credit risk, causes a 1.5% increase in the probability of receiving any wages during the year. We note that our preferred estimates are the reduced form estimates shown in Table 3.

[26]In the Online Appendix Table IAIII we present the results of specifications with alternative transformations of the dependent variable: a) using the hyperbolic sine transformation as an alternative to replacing zeros in the logarithm, and b) using the level of wages.

in Table 2, the coefficient in Column 1 is equal to 6.7% standard deviations of the log(wages+1) variable. This effect, which implies a higher wage of approximately $480, combines the extensive margin effect driven by a higher probability of receiving any wage income as well as an intensive margin effect of higher salaries conditional on employment. We estimate in a back-of-the-envelope calculation that approximately 53 percent of the higher wage effect is driven by the extensive margin.[27] These calculations imply important effects on both the intensive and extensive margins, which suggest that the labor market is not able to adjust on wages alone. This is consistent with frictions in the labor market (e.g. search frictions or efficiency wages).[28]

Column 2 of Table 4 shows that individuals whose nonpayment flags are publicly available for less time are 1 percentage point less likely to be self-employed. This suggests that individuals appear to use self-employment as a response to unemployment, rather than to invest in an entrepreneurial activity with high growth potential.[29]

Finally, column 3 of Table 4 shows that credit market information affects an individual's total pretax income in a significant manner. That is, individuals' total incomes are higher when their information on past defaults are no longer publicly available. This implies that households are not able to fully offset losses to wage income with income from self-employment activities. The effect of credit information on income appears to be slightly lower in magnitude than the effect on wages. This is consistent with the fact that individuals are able to attenuate part of the effect of credit information on employment through low-return self-employment.

As a robustness test, in the Online Appendix Table IAIV we present the results of running our main regression test on a sample where we shift the definition of Treatment and Control groups one year ahead. That is, we define a Placebo Treated group as individuals who defaulted in 2001 and a Placebo Control group as individuals who defaulted in 2002, and use *employed*, a dummy for positive wage income, and the log of wages plus 1 as outcomes. In all three cases, the estimated coefficient of

[27]We obtain this fraction as follows. First, the average wage of individuals who transitioned from zero wages to positive wage income in event time 2, the year before the past default flag is removed, is 71,200 SEK. Thus, a 3% extensive margin effect from Column 4 in Table 3 corresponds to a wage effect of 2,129 SEK. From the appendix, the coefficient on the regression using wage as the outcome variable implies a total effect of 3,987 SEK (See Column 2 in Table IAIII in the Internet Appendix). Thus, the extensive margin represents a $\frac{2{,}129}{3{,}987} = 53.4\%$ of the total wage effect.

[28]The typically high level of unionization in Sweden contributes to a limited scope for adjustment along the wage margin. For statistics on the trade union density in Sweden see for example https://stats.oecd.org/Index.aspx?DataSetCode=UN_DEN.

[29]See Banerjee, Breza, Duflo, and Kinnan (2015) for an application of this idea in India.

interest is not significantly different from zero at conventional levels and takes the opposite sign.

3.3. **Results by treatment intensity.** Our identification strategy relies on variation in the retention times of nonpayment information induced by the policy change. The regression tests so far show that individuals who were exposed to a shorter retention time have a higher probability of being employed than those who were not exposed to it. To further support our identification, we study whether individuals who were *differentially* exposed to the longer retention times, measured by the time of the year in which they defaulted, experience differential labor market responses.

We proceed by categorizing individuals in our sample who were exposed to shorter retention times in five groups according to the bimonthly period in which they defaulted: February-March, April-May, June-July, August-September, and October-November. This categorization of default cohorts induces a monotonic ordering of exposure to the policy change, defined as the average reduction in the number of months during which the nonpayment flag was available in the credit bureaus, for Treated relative to Control group individuals: the August-September cohort has a one month average reduction, June-July has a three month average reduction, April-May has a five month average reduction, and February-March has a seven month average reduction. If information about nonpayments affects the probability of being employed, we hypothesize that the measure of months of exposure to the policy, i.e. the number of fewer months in which past arrears are reported, should be positively correlated with the probability of being employed during a given year. Note that the October-November cohort has, by construction, a zero month reduction in retention time.

To test this hypothesis, we modify regression model (2.1) by changing the interaction variable $early_i$, which divided individuals into early- and late-in-the-year defaulters, with a set of fixed effects for $exposuremonths_i$, which takes values 1, 3, 5, or 7. Thus, we estimate the following specification:

$$\begin{aligned} 1\left(wages > 0\right)_{i,t} \quad = \quad & \omega_i + \omega_\tau + \omega_t + \sum_{t=1,3,5,7} \beta_t 1\left(exposuremonths_i = t\right) \times treated_i \times post_{i,t} + \\ & \delta \times post_{i,t} + \gamma treated_i \times post_{i,t} + \\ & \sum_{t=1,3,5,7} \lambda_t 1\left(exposuremonths_i = t\right) \times post_{i,t} + \varepsilon_{i,t}. \end{aligned} \tag{3.1}$$

The excluded category of $exposuremonths_i$ corresponds to individuals who defaulted in November-December, who have zero months of exposure to the policy. We run this regression using the dummy $1\left(wages>0\right)_{i,t}$ as the outcome, and limit the post period to the year during which the nonpayment flag is removed. Figure 4 shows a plot of the regression coefficients β_t and the associated 95 percent confidence intervals. Consistent with our identification assumption, the measured effect is stronger for individuals who experienced greater reductions in retention times because of the month in which their default occurred. Further, the pattern is monotonic for three, five, and seven months of exposure. The pattern is very similar for $log(wages+1)$, also shown in the lower panel of Figure 4. One month of exposure corresponds to a increase of 0.14 log wage points, while seven months of exposure corresponds to an increase of 0.28 log wage points.

In Table 5, we allow the treatment effect to be linear in the length of exposure to negative credit information, according to the following specification:

$$\begin{aligned} 1\left(wages>0\right)_{i,t} &= \omega_i+\omega_t+\omega_\tau+\beta exposuremonths_i\times treated_i\times post_{i,t}+ \\ &\quad \delta\times post_{i,t}+\gamma treated_i\times post_{i,t}+ \\ &\quad \sum_{t=1,3,5,7}\lambda_t 1\left(exposuremonths_i=t\right)\times post_{i,t}+\varepsilon_{i,t}. \end{aligned} \tag{3.2}$$

Consistent with the results presented in Figure 4, we find in columns 1 and 2 that one shorter month of retention time is associated with a 0.5 percentage points and 0.6 percentage points increase in the probability of earning positive wage income in the same year and in the first two years after information is deleted, respectively. Similarly, in columns 3 and 4, we show that an additional month of exposure to the negative information causes an increase in $log(wages+1)$ of 0.036 and 0.04, again for the first year or first two years after information is deleted, respectively. Both coefficients are statistically significant at the 1 percent level. We believe that the results in both Figure 4 and Table 5 are consistent with and provide credibility for the identification assumption.

3.4. **Other Results: Mobility and Education.** We explore two additional margins that may be affected by changes in credit market information. First, we measure whether increased retention time affects an individual's geographic mobility. Because landlords commonly check a prospective lessee's credit history before signing a lease agreement, we hypothesize that individuals may be more able to move if negative information is held by the credit bureau for a shorter period. Moreover, improved access to employment opportunities may also induce mobility. We test this hypothesis in columns 1 and 2 of Table 6 and define the outcome variable $relocates_{i,t}$ as an indicator for whether an individual moved to a different municipality between years t and $t-1$. In Column 1, we consider the treatment effect for the entire analysis sample and find that individuals who experienced a shorter retention time are 1.1 percentage points more likely to move, relative to a baseline mean of 7.7 percent. While a large effect in relative magnitude, the coefficient is not statistically significant at standard levels (p-value = 0.19). Given that members of our sample have very low home ownership rates (9.6 percent) and that credit checks for residential rental leases are common in Sweden, in column 2, we restrict the sample to the set of individuals who did not own a home in the preperiod. Here we find that individuals who are not home owners are 1.6 percentage points more likely to move across postal codes when their negative credit market information is available to the credit market for less time. While the results are only significant at the 10 percent level, we find them highly suggestive of a type of mobility lock in the rental market because of credit market information.[30]

The results on mobility are unable quantitatively to explain the full employment effect shown in Table 3. Although improvements in mobility to better labor markets induced by the removal of bad credit information may have a causal role explaining the employment results, we perform a bounding exercise and find that this can explain at most 27 percent of the baseline effect of information on employment in Table 3.[31] Again, the direction of causality may also flow the opposite direction – a change in employment status may facilitate relocation.

Second, we ask whether some individuals respond to decreased labor market opportunities by adjusting their demand for additional schooling. When wage jobs become more scarce, the opportunity cost of schooling decreases, which may in turn increase the demand for schooling.[32] This may be especially true in Sweden, where educational loans do not require credit checks and where the costs of education are relatively low. In column 3 of Table 8, we find evidence that education is indeed one margin of

[30] This is similar to the housing lock-in documented by Struyven (2014) in the case of Dutch home-owners with high loan-to-value ratios.

[31] We estimate this fraction as follows. We repeat the mobility regression result conditioning on individuals who moved who also changed employment status, which implies a coefficient of 0.8%. If we fully attribute this coefficient to the causal effect of increased mobility following the early removal of credit information, then mobility can explain up to $\frac{0.8\%}{3\%} = 27\%$ of the baseline effect on employment (denominator from Column 4 in Table 3). Additionally, in unreported results we find that the chronically unemployed and the highly educated, two populations with no measured effect of credit information on employment, are significantly more likely to move when their information on past defaults is removed. This also suggests changes in mobility do not explain our employment results.

[32] See Charles, Hurst, and Notowidigdo (2015) for evidence of this idea in the US.

adjustment used by individuals. Decreased retention time decreases the number of years of education by 0.0355. While the effect is small in magnitude, it is significant at the 5 percent level.

Taken together, our results provide a consistent characterization of the effects of credit market information on labor markets. We interpret these results as the inverse of our treatment effects: information on past defaults reduces the probability that an individual is and remains employed. Individuals respond to this decrease in employment opportunities by turning to self-employment activities and seeking additional education. As a result, individuals earn lower wages and lower total incomes two years after the information is removed from the control group's credit bureau records.

3.5. **Incidence.** It is natural to next ask, for which types of individuals are the employment effects of negative credit information the strongest? First, we explore whether the effects differ by the employment history, namely the preperiod (event time 2) employment status. There is reason to believe that both the previously employed and previously unemployed may experience negative impacts. For the previously employed, there are two mechanisms that may result in negative impacts of credit information on labor market outcomes. First, many individuals in our specific sample are likely underemployed or employed in temporary jobs. Negative credit information may keep any such workers from finding a better or new job for all of the reasons discussed in the Introduction. Second, the condition of being financially constrained, which is also caused by the negative credit information may have a direct impact on worker productivity, earnings, and job tenure. Recall that this mechanism is one reason that employers may choose to screen applicants using credit scores in the first place.

On first thought, one might hypothesize that the effects of increased retention time should be more pronounced for the previously unemployed, who may be more likely to be actively searching for a new job. However, there are countervailing factors. For example, individuals with long unemployment spells may already be severely handicapped in the labor market (e.g.., Kroft, Lange, and Notowidigdo (2013)) and may have even stopped actively searching. Thus, the additional impact of negative credit market information may be muted for this group.

In Table 7 we investigate whether the effects of the shorter retention time are stronger for those individuals who were unemployed in the preperiod versus those who were employed. In Column 1 we run our main specification (Equation 2.1) restricting to those employed at event time 2 (i.e., the year before arrear removal). In column 2, we run the same specification but restrict to those without employment in event time 2. We find the effects on both the likelihood of having wage employment and log wages to be similar for both groups.

To further analyze the group of previously unemployed individuals, we explore not only the event time 2 employment status, but the length of the unemployment spell. We define the chronically unemployed to be those without employment at event time 2, and additionally who worked at most one year in the three preperiod years. The nonchronically unemployed are those who are unemployed at event time 2 who reported any amount of work during more than one year in the three preperiod years. We present our main specification restricted to the chronically unemployed in column 3 and to the non-chronically unemployed in column 4. While the regressions suffer from a lack of power, the patterns are nonetheless striking. We find that the effects are much smaller in magnitude for the chronically unemployed both in terms of participating in wage labor and log wages. In turn, the effects are largest in magnitude among all groups for the non-chronically unemployed.

Second, we study how our effects vary for individuals with different levels of education. This variable is correlated with income and is likely to proxy for employment opportunities in general. In Table 8 we present the output of our main regression test for two sub-samples: individuals with 11 or fewer years of completed schooling (the median number of years of schooling), and individuals with more than 11 years of schooling. Columns 1 and 2 show that a shorter retention time strongly increases the probability of employment for individuals with little education, but it has almost no effect on individuals with many years of schooling. Columns 3 and 4 show that this pattern is repeated using log(wage+1) as outcome.

Our results suggest that the impact of negative credit information on employment is felt more acutely by those with lower levels of education. One interpretation of this heterogeneity is that past credit information is one of many signals used by employers to infer an individual's unobserved productivity. For well-educated individuals, this information may be less relevant than other types of information and as such it may be down weighted when the employer tries to infer unobserved productivity. Individuals with little formal education may also have fewer ways to signal their type. As a result, their past defaults may be a stronger signal of future productivity. Additionally, individuals with less education may be relatively more affected by a restriction to their supply of credit. If so, the removal of negative credit information allows them to make investments that result in higher levels of observed employment.[33]

Third, we study whether the effects of the removal of credit information on employment vary depending on local labor market tightness. For this we estimate the average local unemployment rates during 2003 and 2004 at the kommun level, a Swedish geographic division with at least 5,000 inhabitants (there are 290 kommuns in Sweden with, on average, 33,693 inhabitants). In Table 9 we present our regression results, run in separate samples depending on whether the individual's kommun of residence had a preperiod level of unemployment higher or lower than 3.85 percent (the cross sectional median). Columns 1 and 2 show that the employment effect of shorter retention time, measured with the positive wages dummy, is only concentrated in areas with low unemployment.[34] Columns 4 through 6 show that wages follow the same pattern. This evidence suggests that credit information has stronger effects in areas with low unemployment. Thus, the employment cost of default is more severe for bad individuals in "good times" (i.e., low unemployment areas) than for households exposed to systematic shocks.

4. Additional evidence

We have documented economically large employment costs of default. These costs are borne disproportionally amongst individuals who are previously employed or who are unemployed but looking for a job, amongst individuals with little formal education, and in areas of Sweden with relatively low unemployment. This large employment cost of default is at least partly induced by the direct effect of information on access to credit, which restricts individuals ability to search for a job as well as to invest in becoming more productive. Here we present some additional evidence that suggests that information itself, through for example employer screening, may be behind some of the results we document.

We start in Table (10) by studying whether the removal of negative credit information affects the affected individual's spouse's employment. Intuitively, if households are restricted in their access to credit, then a relaxation of credit constraints would also allow an individual's spouse to supply more labor. At the margin, this would result in more employment for both the individual and the individual's spouse. In that case, the coefficient of interest of our main test using a measure of the spouse's

[33]In unreported results we find, however, that the credit effects are similar for both groups.

[34]In column 3 we exclude the kommun of Stockholm from the sample of low unemployment areas and results are unchanged.

employment as the outcome variable would be positive, that is, would have the same sign as the individual's employment outcome in our main tests.

Although we cannot observe the spouse's employment directly, for each individual in our sample we observe measures of "household disposable income" and "individual disposable income". At the household and individual levels, disposable income is calculated by our data provider by adding up all income sources and subtracting allowances for dependents (children) and adjusting for the cost of living in a particular area. From these measures, we construct the spouse's disposable income by subtracting the individual's disposable income from the household's disposable income.[35]

In columns 1-3 of Table (10) we present the output of regression (2.1) using as outcomes the individual's disposable income, the household total disposable income, and the spouse's disposable income, respectively. The spouse's disposable income can be negative due to government transfers and adjustments, which makes it impossible to use a logarithm plus one approach like we do in previous sections for other continuous outcome variables.[36] We restrict the sample of individuals to those that appear as non-single as of event time 2, the year before past defaults are removed from the individual's credit record.[37] Although underpowered, these tests show that the individual's and household's disposable income increases when their information on past defaults is removed.[38] However, column 3 shows that the spouse's disposable income does not vary in a statistically significant manner with negative credit information, and, if anything, the point estimate is negative. This evidence suggests that access to credit, brought about through better information, does not necessarily help relax household-level credit constraints that prevent access to labor markets. This non-result is perhaps even more surprising given that the credit information of spouses is likely correlated.[39]

[35]We winsorize each of these these variables at the 99th percentile.

[36]These specifications using levels are comparable to the one we present in the Online Appendix Table IAIII using wage as the outcome.

[37]The restricted sample includes households categorized as husband-wife and unmarried cohabitation ("sambo"). We also include single father and single mother households where the youngest child at home is 18 years old or above. We exclude households categorized as single, and single father and mother households with at least one child under the age of 18 at home.

[38]For comparability with our previous results, we present estimates using the logarithm of individual and household disposable income plus one on columns 4 and 5 of Table (10) and note strongly significant effects of the removal of past of defaults on these outcomes, consistent with the evidence in the previous section.

[39]Thus, it is possible that the spouse actually increases labor supply when the individual is unable to find a job due to negative credit information (Blundell, Pistaferri, and Saporta-Eksten (2016)).

Second, in Table 4 we document that individuals whose credit information is retained for fewer months are less likely to be self-employed. It is plausible that investments required for self-employment are likely to be of the same order of magnitude (if not larger) of any investments needed to enter labor markets and supply labor. As a result, individuals with a bad credit record appear to be unconstrained to pay the fixed cost required to become self-employed, which makes it less likely that they are at the same time constrained to pay a fixed cost to enter the labor market.

Third, we can measure credit record inquiries, i.e. credit checks, by non-financial institutions. Based on information supplied by the credit bureau, these institutions include utility companies, legal representation, and crucially, employers.[40] That is, anecdotally, employers do ask for credit checks of potential employees. If individuals increase their supply of labor when information on past defaults is erased, we expect credit report inquiries of non-financial institutions to increase for Treated individuals who defaulted early in the year relative to Control, late-in-the-year defaulters, after their past default is removed. Thus, we hypothesize that the coefficient of interest (the triple interaction) of regression (2.1) using the number of non-financial credit checks as the outcome should be positive and significant. Column 1 in Table 11 shows that this is not the case: the number of inquiries is not causally affected by the shorter retention time of information on past defaults. On the other hand, column 2 suggests that financial inquiries are significantly increased in the presence of shorter retention times of information. This is consistent with the idea that credit card companies and other lenders are actively pursuing individuals whose flag of past nonpayment is removed from the credit bureau. However, job search behavior does not appear to change.

Fourth, we return to our estimates of the employment effects of the removal of credit information across regions with different unemployment rates shown in Table 9. As noted above, the table shows strong effects in areas with low unemployment, but the effects of the removal of credit information are not detectable in areas with high unemployment. However, columns 7, 8, and 9 of Table 9 show that individuals in areas with high and low unemployment experience an increase in their credit limit following early removal of their credit information. This asymmetric response of employment and credit outcomes to the removal of information again suggests that relieving credit constraints alone is insufficient to generate an effect on employment. Indeed, upon removal of credit information, credit constraints are lifted in both high

[40]We cannot, unfortunately, determine precisely which of these non-financial institutions is performing the checks in our sample.

and low unemployment areas, but only in the latter does this have a differential effect on the probability of being employed.

Moreover, in their analysis of duration dependence on firm hiring decisions, Kroft, Lange, and Notowidigdo (2013) perform a similar heterogeneity analysis to argue that their results are more consistent with employer screening, as modeled in Lockwood (1991), than with skill depreciation. They argue that the average quality of the unemployed is lower in tight labor markets than in labor markets with more unemployment, and thus screening is likely to be more valuable. In our setting, screening on credit information is also likely to be more valuable in tight markets.

We conclude this section by highlighting that the employment effects we measure are likely to be a combination of increased search efforts and increased investment in productivity following a relaxation of credit constraints, but that these credit related effects are likely combined with the fact that employers screen potential employees using credit information. This is compounded by the fact that in Sweden, many of the small investments that households would be required to make in order to work for a wage, including health care and child care, are already covered by a social safety net. This suggests that the credit supply mechanism is likely to be diminished in this particular setting.

5. Conclusion

We combine a unique natural experiment in Sweden with detailed credit and labor market data to document that credit market information has economically important effects that spill over onto other domains of a borrower's life, namely the borrower's success in the labor market. In particular, we find robust evidence that an earlier deletion of negative credit information makes individuals more likely to be employed, and as a result, they earn higher incomes. These results highlight an understudied interlinkage between credit and labor markets.

We also show that when labor market opportunities become scarce, individuals seek out self employment and schooling as alternatives. These results indicate that for our sample of low income Swedes, self-employment appears to be an inferior alternative to the wage labor market. This finding resonates with the narrative in the entrepreneurship literature that many businesses owned by low income groups are not primed for transformative growth. The schooling response to the unemployment caused by negative credit information is also consistent with prior literature.

Our results suggest that the costs of default through information sharing in credit bureaus have profound effects on the livelihoods of individuals. Ex-ante, such effects strengthen the incentive mechanism of the credit bureau. For example, individuals may want to continue to service underwater mortgages if the labor market costs are sufficiently high.[41] However, a temporary shock causing an individual to default may have lasting and profound consequences. These results also imply that it may be very difficult for households to use their labor supply to smooth consumption when their credit record is poor. Furthermore, damage from credit information errors may be amplified through the labor market channel.[42]

[41]Extrapolating to a different market and context, labor market costs may help to explain why strategic default was not even more common during the housing crisis (Foote, Gerardi, and Willen (2008b)).

[42]For example, see http://www.forbes.com/sites/halahtouryalai/2013/12/17/should-your-credit-score-matter-on-job-interviews-senator-warren-says-no-aims-to-ban-employer-credit-checks/.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, 2015, House prices, collateral, and self-employment, *Journal of Financial Economics* 117, 288 – 306.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel, 2015, Do banks pass through credit expansions? the marginal profitability of consumer lending during the great recession, Discussion paper National Bureau of Economic Research.

Autor, David, and David Scarborough, 2008, Does job testing harm minority workers? evidence from retail establishments, *Quarterly Journal of Economics.*

Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnan, 2015, Do credit credit constraints limit entrepreneurship? heterogeneity in the returns to microfinance, *Working Paper.*

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten, 2016, Consumption inequality and family labor supply, *American Economic Review* 106, 387–435.

Bos, Marieke, Susan Carter, and Paige Marta Skiba, 2012, The pawn industry and its customers: The united states and europe, *Vanderbilt Law and Economics Research Paper.*

Bos, Marieke, and Leonard I Nakamura, 2014, Should defaults be forgotten? evidence from variation in removal of negative consumer credit information, *Federal Reserve Bank of Philadelphia Working Paper.*

Brown, Martin, and Christian Zehnder, 2007, Credit reporting, relationship banking, and loan repayment, *Journal of Money, Credit and Banking* 39, 1883–1918.

Charles, Kerwin Kofi, Erik Hurst, and Matthew J Notowidigdo, 2015, Housing booms and busts, labor market opportunities, and college attendance, .

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull, 2007, A quantitative theory of unsecured consumer credit with risk of default, *Econometrica* 75, 1525–1589.

Chatterji, Aaron K, and Robert C Seamans, 2012, Entrepreneurial finance, credit cards, and race, *Journal of Financial Economics* 106, 182–195.

De Janvry, Alain, Craig McIntosh, and Elisabeth Sadoulet, 2010, The supply-and demand-side impacts of credit market information, *Journal of development Economics* 93, 173–188.

Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz, 2016, The value of postsecondary credentials in the labor market: An experimental study, *American Economic Review* forthcoming.

Djankov, Simeon, Caralee McLiesh, and Andrei Shleifer, 2007, Private credit in 129 countries, *Journal of Financial Economics* 84, 299–329.

Dobbie, Will, and Jae Song, 2015, The impact of loan modifications on repayment, bankruptcy, and labor supply: Evidence from a randomized experiment, *Working Paper*.

Einav, Liran, and Jonathan D Levin, 2013, The data revolution and economic analysis, .

Elul, Ronel, and Piero Gottardi, 2015, Bankruptcy: Is it enough to forgive or must we also forget?, *American Economic Journal: Microeconomics* forthcoming.

Foote, Christopher L, Kristopher Gerardi, and Paul S Willen, 2008a, Negative equity and foreclosure: Theory and evidence, *Journal of Urban Economics* 64, 234–245.

——— , 2008b, Negative equity and foreclosure: Theory and evidence, *Journal of Urban Economics* 64, 234–245.

Gerardi, Kristopher, Kyle F Herkenhoff, Lee E Ohanian, and Paul Willen, 2013, Unemployment, negative equity, and strategic default, *Working Paper*.

González-Uribe, Juanita, and Daniel Osorio, 2014, Information sharing and credit outcomes: Evidence from a natural experiment, *Working Paper*.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen, 2014, Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and 'normal' economic times, *NBER Working Paper*.

Herkenhoff, Kyle F, 2013, The impact of consumer credit access on unemployment, *mimeo*.

——— , and Gordon Phillips, 2015, How credit constraints impact job finding rates, sorting & aggregate output, *Working Paper*.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, 2014, Can unemployment insurance spur entrepreneurial activity?, *NBER Working Paper*.

Jayachandran, Seema, 2006, Selling labor low: Wage responses to productivity shocks in developing countries, *Journal of Political Economy* 114, 538–575.

Karlan, Dean, and Jonathan Zinman, 2009, Expanding credit access: Using randomized supply decisions to estimate the impacts, *Review of Financial studies* p. hhp092.

Kehoe, Patrick, Virgiliu Midrigan, and Elena Pastorino, 2014, Debt constraints and employment, *Working Paper*.

Kroft, Kory, Fabian Lange, and Matthew J Notowidigdo, 2013, Duration dependence and labor market conditions: Evidence from a field experiment, *The Quarterly*

Journal of Economics 128, 1123–1167.

——— , and Lawrence F Katz, 2015, Long-term unemployment and the great recession: the role of composition, duration dependence, and non-participation, *Journal of Labor Economics* forthcoming.

Liberman, Andres, 2015, The value of a good credit reputation: Evidence from credit card renegotiations, *Journal of Financial Economics* forthcoming.

Livshits, Igor, James MacGee, and Michï¿œle Tertilt, 2007, Consumer bankruptcy: A fresh start, *American Economic Review* 97, 402–418.

Lockwood, Ben, 1991, Information externalities in the labour market and the duration of unemployment, *The Review of Economic Studies* 58, 733–753.

Low, Hamish W., 2005, Self-insurance in a life-cycle model of labour supply and savings, *Review of Economic Dynamics* 8, 945 – 975.

Mian, Atif, and Amir Sufi, 2010, The great recession: Lessons from microeconomic data, *The American Economic Review* pp. 51–56.

Miller, Margaret J., 2000, Credit reporting systems around the globe: the state of the art in public and private credit registries, *Credit reporting systems and the international economy. Cambridge, MA: MIT Press.*

Mullainathan, Sendhil, and Eldar Shafir, 2013, *Scarcity: Why having too little means so much* (Macmillan).

Musto, David K, 2004, What happens when information leaves a market? evidence from postbankruptcy consumers, *The Journal of Business* 77, 725–748.

New York Attorney General, State of, 2015, A.g. schneiderman announces groundbreaking consumer protection settlement with the trhee national credit reporting agencies, *Press Release, available at http://www.ag.ny.gov/press-release/ag-schneiderman-announces-groundbreaking-consumer-protection-settlement-three-national.*

Padilla, A Jorge, and Marco Pagano, 2000, Sharing default information as a borrower discipline device, *European Economic Review* 44, 1951–1980.

Pagano, Marco, and Tullio Jappelli, 1993, Information sharing in credit markets, *The Journal of Finance* 48, 1693–1718.

Pijoan-Mas, Josep, 2006, Precautionary savings or working longer hours?, *Review of Economic Dynamics* 9, 326–352.

Schmalz, Martin, David Sraer, and David Thesmar, 2015, Housing collateral and entrepreneurship, *Journal of Finance* forthcoming.

Struyven, Daan, 2014, Housing lock: Dutch evidence on the impact of negative home equity on household mobility, *Working Paper*.

Figures

Figure 1. Time line

This figure depicts the timeline of the policy change that enforced a three year retention time for reporting defaults and how this policy generated variation in the retention time of the nonpayment flag for individuals with nonpayments in different moments of the year. In particular, the figure shows that individuals whose nonpayment occurred early 2001 had a reduced retention time of past nonpayments. In contrast, individuals whose nonpayment occurred early in 2000 were reported in the credit registries until October 2003.

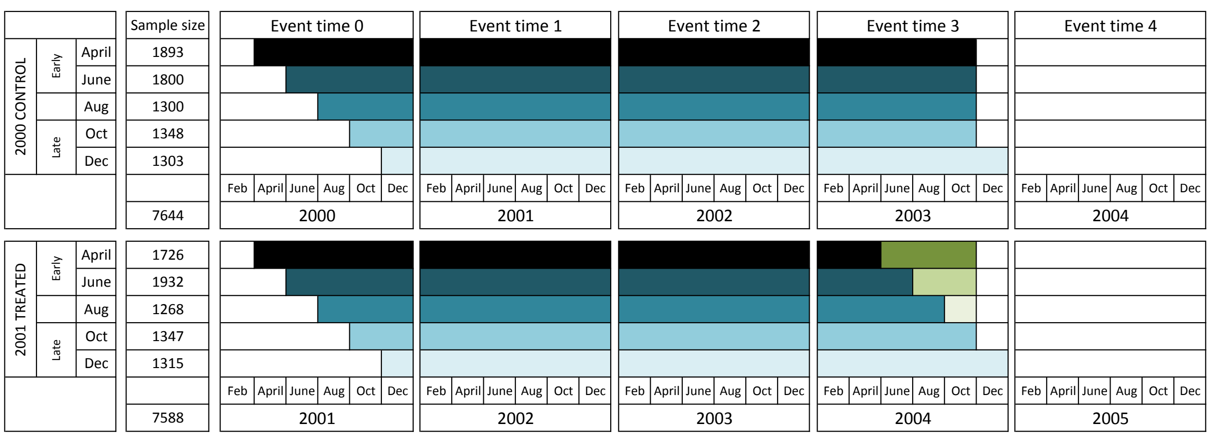


Figure 2. Frequency of removal of nonpayment flag over time
This figure displays the distribution of the removal of nonpayments over time. In the old regime the credit bureau removed all negative arrears that were eligible for removal once a year, on December 31. Each nonpayment was eligible for removal in the third year after the year in which it was received. Because of the bimonthly feature of our data, and because removals are inferred as differences in the stock of reported defaults, these nonpayments corresponds to the February-March bi-month (labeled February). This regime ended at the end of September 2003, when the law change came into effect and the credit bureau was forced to stop reporting all negative flags exactly three years to the day after the default was first reported.

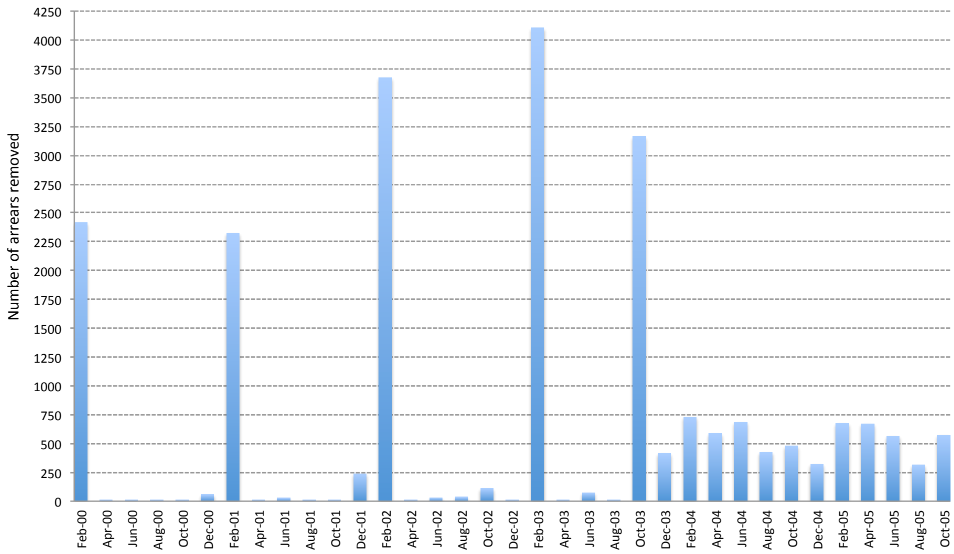


FIGURE 3. Pre-trends

This figure shows that there is no difference in the preperiod trends (before the policy change) of the difference between Early and Late defaulters, in the Treated and Control group for our main outcomes. The top panel shows preperiod trends for *employed* and 1(wages>0), which equals one if an individual received any wage income, the lower panel for log(wages+1) and log(income+1) where zeros have been replaced by 1. The blue lines represent the differences in averages of the respective outcome variables between individuals who defaulted early in the year (high exposure) and individuals who defaulted late in the year (low exposure), for individuals in the Control group. The red line represent the same difference for individuals in the Treated group.

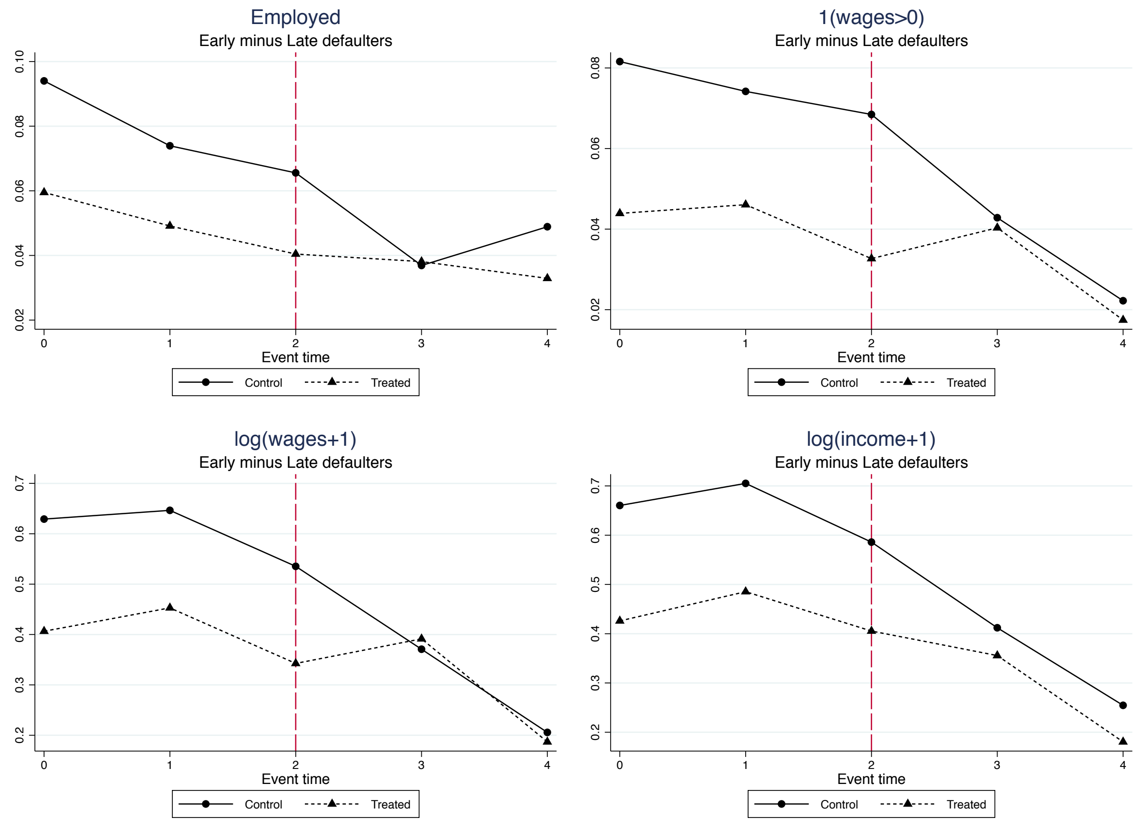
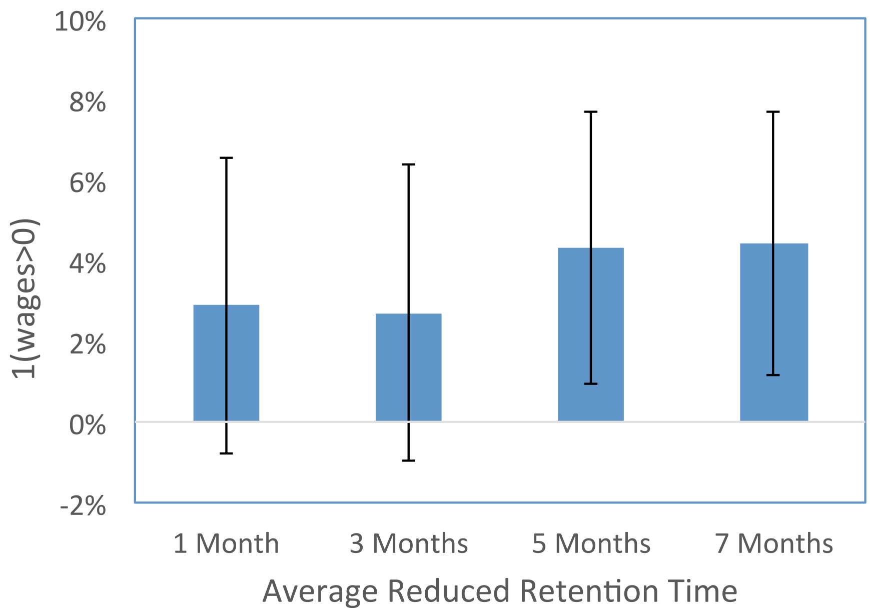
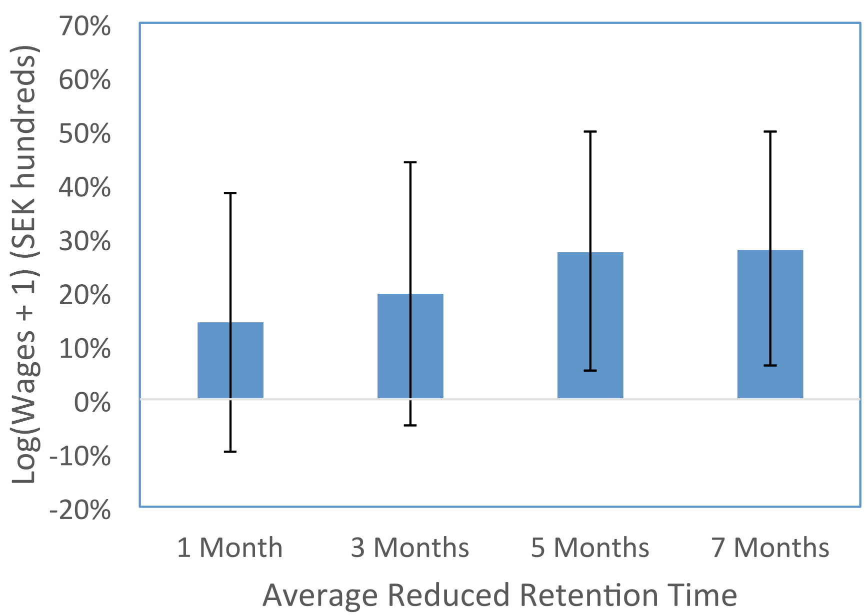


FIGURE 4. Retention time exposure and employment status
This figure depicts that the effect of credit information on labor market outcomes is monotonically stronger with actual exposure to the policy. The graphs show the estimated coefficients of the regression model with varying intensity of exposure (regression (3.1)) versus exposure, defined as the number of months by which individuals who defaulted in the Treated group had their default removed from the credit registry before individuals who defaulted in the Control group. The top panel shows the coefficients using a dummy for positive wages as an outcome, and the lower panel displays the coefficients using the logarithm of wage income as outcome.





Tables

Table 1. Correlation between lagged credit scores and unemployment

This table documents that past credit score (in the case of Sweden, higher score means worse repayment history) is negatively correlated with the probability of being employed using the following OLS regression:

$$1\left(wages_{i,t} > 0\right) = \alpha + \beta_l creditscore_{i,t-l} + \gamma X_i + \varepsilon_{i,t}$$

Employment is defined as a positive wage income, documented by tax records, $1\left(wages_{i,t} > 0\right)$. The table shows the results of a regression of the employment dummy on lags of log credit scores (higher score represents a worse borrower, opposite of FICO scores and other measures used in the US), using a battery of demographic controls (columns 1 and 2) and fixed effects (columns 3 and 4). Controls X_i include gender, age, marital status fixed effects, income, a dummy that equals one for individuals who live in one of Sweden's large cities, and a dummy for past nonpayment flags. Data is a yearly panel of a random sample of the universe of Swedish individuals with a credit score, between 2000 and 2005. Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)	(4)
Dependent variable	$1\left(wages_{i,t} > 0\right)$	$1\left(wages_{i,t} > 0\right)$	$1\left(wages_{i,t} > 0\right)$	$1\left(wages_{i,t} > 0\right)$
$log(creditscore)_{i,t-1}$	-0.0345***		-0.0046***	
	(0.0005)		(0.0009)	
$log(creditscore)_{i,t-2}$		-0.0524***		-0.0023*
		(0.0018)		(0.0011)
Controls	YES	YES		
Ind FE			YES	YES
Obs	70,610	55,635	70,610	55,635
R^2	0.0093	0.0121	0.0013	0.0010
Individuals	15,308	15,060	15,308	15,060

Table 2. Outcome variables and summary statistics

Panel A presents the definition of outcome variables used throughout the paper. Panel B presents selected summary stats as of the three years before nonpayments are deleted, which correspond to 2000 to 2002 for the Treated group (the cohort that defaulted in 2000), and 2001 to 2003 for the Control group (the cohort that defaulted in 2004).

Panel A: variable definition

Dependent variables	
Employed	dummy; one if the individual is employed conditional on being in labor force.
1(wages > 0)	dummy; one if the individual has positive income from work.
log(income + 1)	Log of pretax income, in 100 SEK; zeros replaced by 1.
log(wages + 1)	Log of income from work, in 100 of SEK; zeros replaced by 1.
Self-Employed	dummy; one if the individual received positive wages from entrepreneurship.
Relocates	dummy; equals one if individual's residence is in a different county from previous year.
Years of schooling	Number of years of completed education, inferred from end of year level of education.
Financial inquiries	number of requests for an individuals' credit report by financial institutions.
Nonfinancial inquiries	number of requests for an individuals' credit report by nonfinancial institutions.

Panel B: summary statistics

	(1)	(2)	(3)
Dependent variables	mean	std dev	median
Employed	0.43	0.50	
1(wages > 0)	0.79	0.40	
log(income + 1)	5.62	2.91	7.03
log(wages + 1)	5.57	2.97	7.04
Self-employed	0.05	0.21	
Relocates	0.07	0.27	
Years of schooling	10.70	1.76	11
Financial inquiries	0.52	1.05	
Nonfinancial inquiries	0.54	0.95	
Age	42.83	13.00	42
Male	0.60	0.49	
Home owner	0.09	0.29	
Number of individuals		15,232	

Table 3. Employment outcomes

This table shows that public information on past defaults causally reduces employment. The table shows the coefficient β from regression:

$$\begin{aligned} employed_{i,t} &= \alpha_i + \omega_t + \nu_\tau \beta early_i \times treated_i \times post_{i,t} + \delta \times post_{i,t} \\ &+ \gamma treated_i \times post_{i,t} + \lambda early_i \times post_{i,t} + \varepsilon_{i,t}. \end{aligned}$$

Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	employed	employed	employed	$1\,(wages > 0)$	$1\,(wages > 0)$	$1\,(wages > 0)$
β	0.0280**	0.0203*	0.0125	0.0298**	0.0299***	0.0295**
	(0.013)	(0.012)	(0.014)	(0.012)	(0.011)	(0.014)
Post period	1 year	2 years	only year 2	1 year	2 years	only year 2
Obs	50,623	63,113	50,482	50,623	63,113	50,482
R^2	0.002	0.003	0.003	0.007	0.024	0.027
Individuals	12,664	12,664	12,664	12,664	12,664	12,664

Table 4. Wages, income, and self-employment

This table shows the effects of credit information on (log)wage income, self-employment, and (log)income, using our main regression model:

$$\begin{aligned} outcome_{i,t} &= \alpha_i + \omega_t + \nu_\tau \beta early_i \times treated_i \times post_{i,t} + \delta \times post_{i,t} \\ &+ \gamma treated_i \times post_{i,t} + \lambda early_i \times post_{i,t} + \varepsilon_{i,t}. \end{aligned}$$

Zeros are replaced by one in the log outcomes. Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)
Coefficient	log(wages + 1)	self-employed	log(income + 1)
β	0.1995***	-0.0137**	0.1410*
	(0.077)	(0.005)	(0.075)
Post period	2 years	2 years	2 years
Obs	63,113	63,113	63,113
R^2	0.030	0.003	0.040
Individuals	12,664	12,664	12,664

TABLE 5. Employment outcomes with varying treatment intensity

This table shows the output of a regression that estimates the effect of longer retention time of nonpayment flags on the probability of receiving any wage income during the year. The table shows contains the coefficient β from regression:

$$\begin{aligned} 1\left(wages>0\right)_{i,t} &= \omega_i+\omega_t+\omega_\tau+\beta exposuremonths_i\times treated_i\times post_{i,t}+\\ &\delta\times post_{i,t}+\gamma treated_i\times post_{i,t}+\\ &\sum_{t=1,3,5,7}\lambda_t 1\left(exposuremonths_i=t\right)\times post_{i,t}+\varepsilon_{i,t}.. \end{aligned}$$

There are 15,232 individuals in this sample instead of 12,664 as in previous tables because we include the June-July cohort of defaulters, which is not included in the previous tests to balance individuals with high and low exposure to the longer retention time. Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)	(4)
Coefficient	$1\left(wages>0\right)$	$1\left(wages>0\right)$	log(wages + 1)	log(wages + 1)
β	0.0051**	0.0059***	0.0364***	0.0398***
	(0.002)	(0.002)	(0.013)	(0.013)
Post period	1 year	2 years	1 year	2 years
Obs	60,891	75,911	60,891	75,911
R^2	0.007	0.024	0.018	0.030
Individuals	15,232	15,232	15,232	15,232

Table 6. Additional results: mobility and education

This table demonstrates effects of credit market information on household mobility and education. The table contains the coefficients and standard errors for our linear triple difference in difference estimations, using relocates, which is a dummy that equals one if a individual's residence is in a different county and not missing from the previous event time year, and "years of schooling", which measures the number of years of education as per the individual's last completed level of education as outcomes. The number of observations is lower for "relocates" as it is defined in differences from the previous event time year, so sample period only includes event times 1 through 4 (drops event time 0). Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)
Coefficient	relocates	relocates	years of schooling
β	0.0118	0.0159*	-0.0355**
	(0.009)	(0.009)	(0.014)
Post period	2 years	2 years	2 years
Sample (at event time 2)	full	non-homeowners	full
Obs	50,229	45,356	60,313
R^2	0.001	0.001	0.015
Individuals	12,664	11,441	12,414

Table 7. Heterogeneity by preperiod employment history

This table shows differential effects of credit information on employment depending on preperiod employment status. The table shows the regression output of our main regression model (2.1) for different sub-samples. In both panels A and B, column 1 restricts to a sample of individuals who are employed ($employed_{i,t}$=1) as of event time 2, the year before their information on nonpayments is removed. Column 2 restricts the sample to individuals who are unemployed as of event time 2. columns 3 and 4 split the sample of unemployed individuals. Column 3 restricts the sample to individuals who are chronically unemployed as of event time 2, defined as those individuals who have been unemployed for 2 or more years in the 3 year preperiod. Column 4 restricts to unemployed individuals who are not chronically unemployed. Panel A uses a dummy for positive wage income as outcome. Panel B uses log(wage+1), where zeros have been replaced by 1, as outcome. The post period includes 2 years after information is deleted (event times 3 and 4). Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

Panel A

	(1)	(2)	(3)	(4)
Coefficient	$1\,(wages > 0)$	$1\,(wages > 0)$	$1\,(wages > 0)$	$1\,(wages > 0)$
β	0.0336**	0.0319*	0.0196	0.0578*
	(0.014)	(0.016)	(0.019)	(0.030)
Post period	2 years	2 years	2 years	2 years
Sample (at event time 2)	employed	unemployed	unemployed: chronic	unemployed: nonchronic
Obs	27,114	34,682	24,071	10,611
R^2	0.050	0.016	0.009	0.065
Individuals	5,424	6,942	4,819	2,123

Panel B

	(1)	(2)	(3)	(4)
Coefficient	log(wages + 1)	log(wages + 1)	log(wages + 1)	log(wages + 1)
β	0.2704**	0.1970*	0.0761	0.4505**
	(0.109)	(0.107)	(0.124)	(0.202)
Post period	2 years	2 years	2 years	2 years
Sample (at event time 2)	employed	unemployed	unemployed: chronic	unemployed: nonchronic
Obs	27,114	34,682	24,071	10,611
R^2	0.072	0.018	0.014	0.067
Individuals	5,424	6,942	4,819	2,123

Table 8. Heterogeneity by preperiod education levels

This table shows differential effects of credit information on employment depending on preperiod level of education. The table shows the regression output of our main regression model (2.1) for different sub-samples: individuals with 11 or less completed years of schooling, and individuals with more than 11 years of schooling. Outcomes are positive wage income and log(wages+1), where zeros have been replaced by 1, as defined previously. The post period includes 2 years after information is deleted (event times 3 and 4). Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)	(4)
Coefficient	$1\,(wages > 0)$	$1\,(wages > 0)$	$log(wages + 1)$	$log(wages + 1)$
β	0.0440***	-0.0003	0.2982***	0.0102
	(0.013)	(0.021)	(0.091)	(0.147)
Post period	2 years	2 years	2 years	2 years
Sample (at event time 2)	$\leq$ 11years	>11 years	$\leq$ 11years	>11 years
Obs	44,543	16,240	44,543	16,240
R^2	0.022	0.042	0.029	0.051
Individuals	8,914	3,249	8,914	3,249

Table 9. Effects by labor market tightness

This table shows differential effects of credit information on employment by the local unemployment rate by kommun. The table shows the regression output of our main regression model (2.1) for different sub-samples. Column 1 restricts the sample to communities where the unemployment rate is higher or equal than the cross sectional median of the average in in 2003-2004 (3.85%), while column 2 restricts the sample to communities where the unemployment rate is lower than the median. Column 3 corresponds to the same sample as column 2, but excluding Stockholm kommun. Outcomes are positive wage income (Panel A) and log(wage+1) (Panel B), where zeros have been replaced by 1, as defined previously. Panel C presents the same regression output using the logarithm of credit line as outcome. The post period includes 2 years after information is deleted (event times 3 and 4). Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)
Coefficient	$\mathbb{1}(wages > 0)$	$\mathbb{1}(wages > 0)$	$\mathbb{1}(wages > 0)$
β	-0.0061	0.0523***	0.0348*
	(0.019)	(0.014)	(0.018)
Post period	2 years	2 years	2 years
Sample (average at event time 3 and 4)	high unemployment	low unemployment	low unemployment ex-Stockholm
Obs	23,419	37,979	20,982
R^2	0.016	0.032	0.030
Individuals	4,697	7,623	4,210

	(4)	(5)	(6)
Coefficient	log(wages + 1)	log(wages + 1)	log(wages + 1)
β	-0.0693	0.3687***	0.2561**
	(0.125)	(0.100)	(0.127)
Post period	2 years	2 years	2 years
Sample (average at event time 3 and 4)	high unemployment	low unemployment	low unemployment ex-Stockholm
Obs	23,419	37,979	20,982
R^2	0.024	0.038	0.035
Individuals	4,697	7,623	4,210

	(7)	(8)	(9)
Coefficient	log(creditline + 1)	log(creditline + 1)	log(creditline + 1)
β	0.4251**	0.5116***	0.5590***
	(0.192)	(0.151)	(0.208)
Post period	2 years	2 years	2 years
Sample (average at event time 3 and 4)	high unemployment	low unemployment	low unemployment ex-Stockholm
Obs	23,419	37,979	20,982
R^2	0.015	0.019	0.019
Individuals	4,697	7,623	4,210

Table 10. Treatment effects on individual's and spouse's disposable income

The table shows the regression output of our main regression model (2.1) using the individual's disposable income (Column 1), the household's disposable income (Column 2), and the spouse's disposable income, calculated as the difference between the household's and individual's disposable income (Column 3). Variables are winsorized at the 99th percentile. In columns 4 and 5 we use the logarithm of the individual's disposable income and the household's disposable income respectively, with zeros replace by one. The sample correspond to all individuals that are not single as of event time 2. The post period includes 2 years after information is deleted (event times 3 and 4). Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)	(4)	(5)
Coefficient	individual disp. inc.	household disp. inc.	spouse disp. inc.	log(individual disp. inc. +1)	log(household disp. inc. +1)
β	37.27*	34.25	-5.64	0.1204*	0.1466*
	(20.462)	(26.307)	(22.802)	(0.068)	(0.085)
Preperiod mean					
Post period	2 years	2 years	2 years	2 years	2 years
Obs	23,154	23,154	23,154	23,154	23,154
R^2	0.026	0.021	0.002	0.003	0.002
Individuals	4,667	4,667	4,667	4,667	4,667

Table 11. Treatment effects on number of credit inquiries: financial and non-financial

The table shows the regression output of our main regression model (2.1) using financial (Column 1) and non-financial inquiries (Column 2) as outcomes. The post period includes 2 years after information is deleted (event times 3 and 4). Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)
Coefficient	non-financial inquiries	financial inquiries
β	0.0035	0.1256**
	(0.030)	(0.057)
Preperiod mean	0.542	0.523
Post period	2 years	2 years
Obs	62,929	62,929
R^2	0.044	0.017
Individuals	12,664	12,664

Supplemental Appendix: For Online Publication Only

A. Supplemental tables and figures

Table IAI. Characteristics of early versus late defaulters

The table shows means of selected variables in the 3 preperiod event time years (0, 1, and 2) as defined in Table 2 for individuals in our sample, by whether they defaulted early (February to May) or late (August to November) in the year. An asterisk denotes that the difference between the two means is significant at the 1% level.

	Early	Late
Employed*	0.46	0.39
1(wages>0)*	0.82	0.76
log(income+1)*	5.83	5.28
Wage income*	5.76	5.26
Self-employed*	0.06	0.05
Relocates*	0.07	0.08
Years of schooling*	10.76	10.61
Financial inquiries	0.53	0.52
Nonfinancial inquiries*	0.52	0.56
Age*	43.30	41.91
Male*	0.59	0.60
Home Owner*	0.11	0.08

Table IAII. Sweden macroeconomic indicators

The table shows selected macroeconomic indicators for Sweden for the sample period. Source: Statistics Sweden.

	1999	2000	2001	2002	2003	2004	2005
GDP growth (annual %)	4.53	4.74	1.56	2.07	2.39	4.32	2.82
Inflation, consumer prices (annual %)	0.45	1.04	2.41	2.16	1.93	0.37	0.45
Unemployment, total (% of total labor force)	7.10	5.80	5.00	5.20	5.80	6.50	7.70

Table IAIII. Alternative specifications of wage outcome

The table shows alternative specifications for our baseline wage regressions shown in Table 4. In particular, we define wages using the inverse hyperbolic sine transformation, which can be interpreted as a percentage change (Column 1), and the level of wages in 100 SEK winsorized at the 99th percentile. Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)
Coefficient	inv. hyp. sine	wages
β	0.2321***	39.88*
	(0.087)	(21.93)
Post period	1 year	1 year
Obs	50,623	50,623
R^2	0.018	0.060
Individuals	12,664	12,664

Table IAIV. Placebo test

This table shows the results of running our main regression test on a placebo sample. Here we define the Placebo Treated group as individuals who defaulted in 2001 and the Placebo Control group as individuals who defaulted in 2002. The coefficient β measures the difference in the outcome for individuals in the Placebo Treated group who defaulted early and late in the year, relative to the same difference for individuals in the Placebo Control group, before and after the deletion of their past nonpayment flag, which occurs on event time 3 (2004 for the Placebo Treated, 2005 for the Placebo Control). The post period includes only one event time year as our sample ends in 2005. Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)
Coefficient	employed	$1\,(wages > 0)$	$log\,(wages + 1)$
β	-0.0080	-0.0038	-0.0708
	(0.012)	(0.013)	(0.090)
Post period	1 year	1 year	1 year
Obs	50,802	50,802	50,802
R^2	0.001	0.025	0.026
Individuals	12,713	12,713	12,713

B. Instrumental Variables Analysis

In order to help interpret the magnitudes of our findings vis a vis the OLS relationship, we estimate a regression of employment on contemporaneous log credit score, instrumenting credit score with early x treated x post. The results are displayed below in Table IAV. Given that the employment data are only available at an annual frequency, we use the average credit score across the year in these specifications.

In column 1, we find that for our analysis sample, the OLS relationship between log credit score and employment is quite strong. An increase in the credit score of 10% is correlated with 0.06% higher likelihood of wage employment. The OLS relationship becomes smaller in column 2 with the inclusion of individual fixed effects, but retains an economically and statistically significant magnitude.

In column 3, we show that the early x treated x post individuals do indeed experience an improvement in their credit score. In fact, their credit score improves (decreases) by 20% in response to the shortening of arrear retention.[43] Finally, in column 4, we show the instrumental variables estimates of the relationship between the log credit score and employment. We find that the IV estimate is substantially larger than the OLS estimate. Taken at face value, the IV suggests that a 10% decrease (improvement) in credit score causes an increase in employment of 1.495%.

This large discrepancy between IV and OLS is likely due to several factors. Importantly, the exclusion restriction is unlikely to be satisfied in this setting for two reasons. First, non-financial companies, including employers, do not observe the credit score when they perform a credit check. Thus they likely respond to the nonpayment flag directly. This channel likely leads to an upward bias of the OLS. Second, it is unclear that even financial institutions use the credit score as a sufficient statistic for credit-worthiness. If prospective lenders use information about arrears in addition to the credit score, then the IV estimates are again likely upward biased. Furthermore, the treatment effects are could be highly non-linear. It is possible that the complier population, which identifies the IV, is quite different from the average population described by the OLS.

Given these reservations in interpreting the IV estimates, we focus on the reduced form results throughout the body of the paper. However, this exercise does suggest

[43] Recall that the Swedish credit score is a default probability. Larger credit scores indicate worse credit records.

that our reduced form magnitudes are quite economically significant in comparison to the OLS.

TABLE IAV. Instrumental variable estimation: the effect of a better credit score on employment

This table shows the results of an IV estimation using $1\left(wages_{it} > 0\right)$ as the outcome. To get a sense of the magnitude of our estimates, in columns 1 and 2 we show the OLS regression of $1\left(wages > 0\right)$ on the logarithm of the individual's contemporaneous yearly average credit score. Column 1 includes demographic controls (dummies for gender, marital status and education level). Column 2 includes individual fixed effects. In columns 3 and 4 we instrument the logarithm of the individual's yearly average credit score, $log\left(average\ score_{it}\right)$, with the triple interaction $early_i \times treated_i \times post_t$, the variable of interest in our main test as shown in Table 3. Column 3 shows the coefficient of the regression of $log\left(average\ score_{it}\right)$ on the instrument, and the IV columns show the coefficient on $log\left(average\ score_{it}\right)$ of the regression of $1\left(wage_{it} > 0\right)$ on instrumented $average\ score_{it}$. *, **, and *** represent 10, 5, and 1 percent significance level, respectively. Standard errors on the OLS and First Stage clustered at the individual level. Standard errors on IV are normal.

	(1)	(2)	(3)	(4)
	OLS	OLS	First stage	IV
Outcome	$1\left(wages > 0\right)$	$1\left(wages > 0\right)$	$log\left(average\ score\right)$	$1\left(wages > 0\right)$
$early \times treated \times post$			-0.1985***	
			(0.027)	
$log\left(average\ score\right)$	-0.0649***	-0.0175***		-0.1495***
	(0.002)	(0.002)		(0.056)
Mean preperiod average score	63.2	63.2	63.2	63.2
Post period	1 year	1 year	1 year	1 year
Controls	YES			
Individual FE		YES		
Obs	49,828	49,828	49,828	49,828
Individuals	12,607	12,607	12,607	12,607