THE LABOR MARKET EFFECTS OF CREDIT MARKET INFORMATION

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

Abstract. This paper documents that credit market information causally affects labor market outcomes. We exploit a detailed dataset that matches credit and labor market outcomes in Sweden and a policy change that provides quasi-experimental variation in the time that information on past defaults remains publicly available. We show that individuals whose past defaults are publicly available for longer are less likely to have a job and are more likely to be self-employed. These individuals also earn lower incomes on average. Our results highlight how credit information may impose non-credit market related costs to individuals emerging from financial distress. Such costs may increase borrower repayment incentives, but may also amplify negative shocks.

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

Keywords: Household Finance, Labor Demand, Credit Information

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" (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 of consumer credit (Djankov, McLiesh, and Shleifer (2007)). Multi-lateral institutions such as the IMF and World Bank urge countries to adopt registries, citing them as a fundamental step

Date: November 2015.

This version: November 2015.

We thank Nathan Hendren, Andrew Hertzberg, Wei Jiang, Daniel Paravisini, and Thomas Philippon, as well as seminar audiences at Berkeley, Columbia, the Federal Reserve Bank of Philadelphia, NYU, Stockholm University, and Sweden Riksbank for helpful comments. Jesper Böjeryd provided excellent research assistance. All errors are our own.

[§]Stockholm University. Email: marieke.bos@sofi.su.se.

[†]Columbia University. Email: ebreza@gsb.columbia.edu.

[‡]New York University. Email: aliberma@stern.nyu.edu.

towards financial development. Unsurprisingly, when countries do have registries, 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 non-credit outcomes, such as employment. While credit information may affect employment indirectly through its effects on credit supply, more direct channels are also possible. Namely, if non-bank actors make decisions based on credit information, then a signal of a past default itself may directly affect non-credit 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. Importantly, there is ample anecdotal evidence that many employers around the world query credit registries when making hiring decisions.[2]

In this paper, we estimate the causal impact of information of past defaults on labor market outcomes. We make use of detailed credit registry data from Sweden that is matched to labor market outcomes and tax records. We exploit a natural experiment that changed the amount of time that records of past delinquencies were retained on consumer credit reports for a subset of individuals. This policy change allows us to disentangle the causal effect of credit market information on labor market outcomes from a correlation driven by omitted variables or reverse causality. For example, individuals with worse labor market prospects are at the same time more likely to default on their debts.

Our empirical analysis is set in Sweden. Swedish law establishes that every non-payment reported to the credit bureaus must be deleted three years after the delinquency occurred. This signal of non-payment is a binary flag, and therefore differs from a more gradual scale of past repayment behavior, for example FICO scores in the US.[3] Instead, this non-payment signal resembles the US bankruptcy flag, which is binary (i.e, in bankruptcy or not) and is deleted 7 years after bankruptcy (Musto

[1]E.g., 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]47% of firms in the U.S check the credit information of their prospective employees according to: http://www.shrm.org/research/surveyfindings/articles/pages/creditbackgroundchecks.aspx.

[3]The credit bureau in Sweden also provides continuous credit scores, but we focus on the binary non-payment indicator in our main analysis. This non-repayment flag is what employers can observe from an individual's credit record and has a discrete adverse impact on an individual's credit score.

(2004)). Importantly, in Sweden, the three year clock for deletion of past non-payments is reset every time a borrower defaults. Therefore, removal of the non-payment flag only happens three years after the defaulter repays and does not incur any subsequent delinquencies.

The natural experiment occurred as follows. Before October 2003, the law governing the removal of arrears was implemented by the credit registry in such a way that information on past non-payments was deleted on the last calendar day (i.e., December 31st) of the third year after the non-payment. Beginning in October 2003, the law was reinterpreted to mean that the information must be deleted exactly three years to the day after it was generated. Importantly for identification, the key impetus for this change was technological, as the upgrade of the computer systems used by the registry improved its data management capabilities. We note that this policy change was first exploited by Bos and Nakamura (2014), who find that shorter retention times result in a restriction of the supply of credit and a higher likelihood of default. 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 not affected by the policy change, and therefore the record of her default was publicly available in the credit bureau until October 2003, three years and eight months later. Next, consider an individual who defaulted in February 2001. She was affected by the policy change, so the record of her default was publicly available in the credit bureau only until Feburary 2004, exactly three years later. Thus, defaulting in February of 2000 or Feburary of 2001 led to different retention times of the non-payment flag, namely an eight-month reduction in arrear retention 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 non-payment flag caused by the policy change to identify the causal effect of credit market information on labor market outcomes. We refer to the 2001 cohort of defaulters as the Treated group, as they were affected by the policy change, and the 2000 cohort of defaulters as the Control group. The policy caused a decrease in the average arrear retention time for members of the Treated group relative to the control group. However, a simple comparison of Treated and Control individuals before and after the removal of 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 size of the retention time reduction due to the policy change varied by the calendar month of default. In our example above, a Treated individual who defaulted in February experienced an eight month reduction 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 had 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 non-payment 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 four (and more) years after default (i.e., post-deletion) relative to the first three years after default (i.e., pre-deletion).[4] Our approach is therefore a triple-differences identification strategy.

We find that a Treated individual who had defaulted early in the calendar year is approximately three percentage points more likely to be employed the year in which their non-payment information is removed from the credit registry. 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., 8 months for the Feburary defaulters, 7 months for the March defaulters, etc.– and the probability of being employed. Further, we find that Treated individuals 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 than Control individuals. 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.

In Section 2 below we discuss in detail how models of labor supply and labor demand could lead to changes in equilibrium employment in response to changes in

[4]In addition, we restrict our sample to those individuals who did not default again in the subsequent 24 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.

credit information. Our results are best explained by mechanisms involving labor demand. If credit information is predictive of job performance, then employers may want to use this information to screen workers. Consistent with this interpretation, we find that the employment effects of credit market information are strongest in tighter labor markets (i.e., those with less unemployment), where the benefits from employee screening are likely to be the highest. Further, we find no detectable effects of the policy change on labor markets with above-median levels of unemployment.

Models of labor supply fit less well with our results. Most channels from improved credit information to labor supply would operate through the increase in credit supply resulting from arrear deletion.[5] However, we find results opposite to what many theories would predict. Namely, if workers have precautionary motives and increase labor to smooth consumption (e.g., Jayachandran (2006)), then we would expect increases in credit supply to be associated with decreases in labor supply. Instead we observe increased employment. Similarly, we find that arrear deletion leads to a decrease in entrepreneurship. Again, most models would predict an increase in entrepreneurship in response to a relaxation of credit constraints. Importantly, we can also rule out a debt overhang model of labor supply given that all individuals in our sample frame, both Treated and Control, defaulted and repaid their debts prior to arrear deletion.

Another alternative is that credit constraints may limit the ability for workers to search for a job (i.e., buying a suit, repairing the car). However, we find that credit supply effects are large even in areas with high unemployment, where we find no employment effect. Further, credit checks by non-financial companies are not changing differentially for early defaulters in the Treated group, suggesting no detectable differences in job search patterns. This evidence, along with our self-employment results, suggests that credit supply can't fully explain the patterns in employment.

Finally, labor supply effects may also arise through changes in mobility. Given that many landlords perform credit checks on potential tenants, negative credit information may contribute to mobility lock. While we do find effects on mobility, we perform a bounding exercise and show that such effects can explain at most one quarter of the employment effect.

Our results suggest that information on previous bad repayment behavior following financial distress may preclude households from being employed, resulting in more financial distress and in a negative spiral that they may be unable to break. Credit

[5]Indeed, Bos and Nakamura (2014) find that access to credit increases substantially when information on past defaults is removed.

information may thus induce a multiplier effect on unemployment. As a result, the labor market effects of credit information may impose large *ex post* utility costs on households. 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. However, the welfare implications of the credit-labor market linkage are ambiguous, as costs borne in the labor market may improve *ex ante* repayment incentives, potentially leading to increased financial access. Further, while longer retention times decrease the labor market opportunities of former defaulters, employers may experience productivity gains by using information about arrears to screen workers and improve matching in labor markets (e.g., Autor and Scarborough (2008)).

Policy makers have recently begun to assess the laws governing the use of credit market 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. Our results provide timely evidence to evaluate the tradeoffs involved in this debate.[6]

The remainder of the paper is organized as follows. Section 2 examines and summarizes theories of the direct and indirect effects of credit market information on employment and labor market outcomes. Section 3 describes the data, setting, and empirical strategy. Section 4 presents the results, Section 5 provides an interpretation of our results, and Section 6 concludes.

2. Motivation

There is a strong observational relationship between loan repayment behavior and employment status. We document this stylized fact in Table 1 using a panel data set of a random sample of 15,862 individuals from the Swedish population, matched to tax records. The table presents the output of an OLS regression of $1\,(wage_{i,t} > 0)$, 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 credit scores, $creditscore_{i,t-l}$ ($l = 1$, $l = 2$), with controls X_i, which include demographic characteristics (Columns

[6]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 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?_r=0.

1 and 2) or individual fixed effects (Columns 3 and 4):

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

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).[7] 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 non-payment in their records (omitted variables). In this paper we study whether at least part of this correlation, additionally, reflects a causal 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.

How can credit market information affect equilibrium labor market outcomes? Here, we discuss possible effects of the availability of credit information on both labor demand and labor supply.

We start by analyzing the potential effects of credit market information on labor demand. Such a channel requires negative credit market information to cause a decrease in the actual or perceived labor productivity of workers. The widespread use of credit reports to screen workers (as discussed in Section 1) suggests this indeed may be the case, though there is scant empirical evidence. Credit market information may be predictive of job performance for two reasons. First, past repayment behavior may be correlated with unobservable fixed heterogeneity in productivity (i.e., worker type). Second, either the conditions that caused the negative credit information in the first place or the financial constraints (i.e., decreased credit supply) resulting directly from the negative information may causally affect an employee's productivity (see, for example, Mullainathan and Shafir (2013)). Under both models, an employer would prefer to hire an employee that has a clean credit record to one with a salient

[7] E.g., see Foote, Gerardi, and Willen (2008), Gerardi, Herkenhoff, Ohanian, and Willen (2013).

non-payment event. In our setting, we hold the past default experiences of individuals constant and measure the impacts of revealing different information about those defaults in the credit bureau. While the worker productivity channel alone may cause decreases in wages and employment in response to negative credit information, allowing employers to use that information amplifies those costs to workers. Presumably, this is why many governments such as the city of New York have begun to regulate such practices. Of course, prohibiting employer credit queries has the countervailing effect of decreasing the efficiency of employer screening. We do not attempt to evaluate that tradeoff here.

There are also many channels through which labor supply may be impacted by credit market information. Most such channels stem from a decrease in credit supply caused by the negative credit information (Bos and Nakamura (2014)).

First, a large literature in macroeconomics suggests that households use labor supply to smooth consumption in the face of incomplete markets (e.g., Pijoan-Mas (2006)). Thus in these models, a relaxation of financial constraints through an increase in credit supply should result in a decrease in labor supply for otherwise constrained individuals (Jayachandran (2006)).

Second, standard models of debt overhang (Myers (1977)) applied to households would tend to predict that financial distress should decrease labor supply. This is because while in or near default, workers do not enjoy the full benefit from their labor market efforts; rather, debtors have the first claim to labor market earnings. However, it is important to note that while the individuals in our analysis sample did default in the past, they also managed to make on-time payments for the subsequent two years. Further, our natural experiment only resulted in a full removal of the arrear flag from an individual's credit report if they had not accumulated any new subsequent arrears. Thus, while our sample is credit constrained, the variation is driven by individuals who are not actively delinquent on any debts. Further, by comparing individuals exposed to long and short retention times, our setting holds debt delinquencies constant. This suggests that the standard debt overhang story is unlikely to play a central role in explaining our results. For debt overhang to have bite our setting, alterations to the standard model are required, such as in Elul and Gottardi (2015).

Third, access to credit may affect the costs of job search. Instead of searching for wage employment, individuals could instead choose to pursue self-employment opportunities, which represent a higher opportunity cost of search for a job. Many studies find that decreased access to credit decreases self-employment income (e.g., Chatterji and Seamans (2012), Adelino, Schoar, and Severino (2015), and Greenstone, Mas, and Nguyen (2014)). This should make it more attractive for workers to supply their labor to the market in response to a reduction in credit supply. In the same spirit, matching models of the job market can explain potential effects of access to credit on the intensive margin (length of search) of the job search (e.g., Herkenhoff (2013)). Here, access to credit increases an individual's outside option, which allows individuals to wait longer for a better match.

Access to credit also may play a role in even allowing an individual to search for a job in the first place (the extensive margin of search), as in Karlan and Zinman (2009). This may be the case if individuals must make an initial investment, such as buying a suit, repairing a car, or setting up child care, to enter into the job market. That is, individuals whose negative information is publicly available for longer may be unable, because of liquidity constraints, to supply their labor at all. We note that because in this model individuals would still like to supply their labor, this explanation is similar to demand-driven channels of the effect of credit information on employment. An important caveat to this interpretation is that Sweden, through its welfare state, provides many of these services at a subsidized or zero cost. Hence, this mechanism may be less important in our setting than in the US, for example, in explaining how access to credit interacts with labor supply and equilibrium employment.

Finally, we note that information may affect employment causally through a lock-in effect of mobility. Indeed, if credit information affects individuals' access to housing, then the removal of the signal of past default may causally allow an individual to reolcate and supply her labor in a better market.

In the next section we turn to our empirical tests to uncover the direction and magnitude of the causal relationship of credit information on labor outcomes. Then, in Section 5 we examine our results and contrast them with the empirical predictions of these alternative mechanisms.

3. Data and identification strategy

Here we describe our empirical setting and baseline identification strategy to uncover the effects of credit market information on labor market outcomes.

3.1. **Setting and policy change.**

Swedish credit bureaus. 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, who jointly own the leading Swedish credit bureau (the "Upplysningscentralen") typically define a borrower to be in default when 90 days past due. However, other entities with access to the credit bureau, like phone companies, exercise discretion as to when a consumer is reported as delinquent.

How does an actual non-payment end up being reported through credit bureaus? After an individual misses a payment on any bank or non-bank claim, recovery is handed to either a private debt collector or to the national enforcement agency. In general, private claims (e.g., bank debts) are first pursued by the private debt collection industry, and if unsuccessful, are handed over to the national enforcement agency. On the other hand, all delinquent government claims are directly handled by the national enforcement agency. Once the delinquent claim is handed over to the national enforcement agency it is officially registered in Kronofogdens' public registry.[8] Any entry for more than $10 is then collected and reported on a daily basis by the credit bureaus.

The policy change. Before October 2003, Swedish law mandated all non-payments to be removed from each individual's credit report after three years. In practice, this meant that the credit bureau removed all non-payments that were registered on the consumer's report for at least three years only once a year, generally on December 31st. In 2003, the Swedish government decided to adjust the law so that every non-payment flag would be removed from the credit bureaus, and thus no longer publicly available, *exactly* three years after the day the non-payment 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.

As shown in Figure 2, the adjustment to the law induced a sharp change in the pattern of removal of non-payment flags by the credit bureaus. The figure plots the number of individuals whose non-payment flags are no longer reported in the credit bureau by bi-month. The figure shows that before 2003, non-payment flags were essentially only removed from the credit registry on the last day of the year. In our bi-monthly data, an individual who had a non-payment flag on December 1st, but had that flag removed on December 31st, is first observed without a non-payment flag in February. Thus, the figure shows that before 2003, the vast majority of arrears were removed only once per year, in December (corresponding to spikes in

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

February 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 non-payments that had occurred between January and October 2000 and that had not yet been deleted from the credit bureau. After October 2003, the frequency becomes more smoothly distributed over the year, in effect following the distribution of non-payments during the year.

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 early, and any 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 available flags of default induced by the 2003 policy change in Sweden to approximate this idealized setting. First, we note that before the policy change, the fact that past defaults were only deleted in December induces variation in the length of time in which defaults that had occured three years earlier were reported. However, it is likely that individuals who default at different times during the year differ in ways that are correlated with labor market outcomes. Further, individuals may have been aware of this feature of credit bureaus and chosen 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.

Instead, the policy change induced *unexpected* variation in the length of time that information was retained in the credit bureaus. Hence, individuals who defaulted 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, the Control group, with those who defaulted in 2001, the Treated group, observing that the average retention time is higher for the 2000 cohort. However, this strategy is also likely to be biased as there may be other differences between defaulters in 2000 and 2001 that may be correlated with labor market outcomes.

Instead, we combine the two empirical strategies – Treated vs. Control cohorts and early vs. 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, with the same difference but for individuals whose default was reported the previous year, 2000. 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 time. Individuals in the Control group who defaulted early in 2000, say in March, were subject to three years + 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 the difference in employment outcomes between individuals in the Control group whose default was reported early and late in the year relative to the same difference for individuals in the Treated group whose default was reported early and late in the year would have remained constant before vs. after arrear removal in the absence of the policy change. In Section 4.4.1 below we provide evidence that is consistent with this assumption.

Finally, note that within the group of Treated group 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 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 4 we provide evidence that is consistent with this intuition.

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

3.2. **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 a pawn loan at least once between 1986 and 2012. This sample is particularly well suited for our analysis. Indeed, these individuals are more likely to face periods of financial distress that lead to the reporting of non-payments on their credit reports compared to 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.[9]

Our data corresponds to a panel at a bi-monthly frequency, with data 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 US, 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 non-payment 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 on 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 includes measures of income such as pre-tax income, wages, and income from self-employment. We defer an analysis of summary statistics of our main outcome variables until after we've presented our sample selection criteria.

3.3. **Implementation of empirical strategy.** We exploit the policy change that reduced the length of time for which non-payments were reported in the Swedish credit registry to identify the causal effect of credit information on labor market outcomes. As we mentioned in the Introduction, while we analyze the same natural experiment as Bos and Nakamura (2014), we follow a different empirical strategy.

In order to isolate those individuals that were most likely affected by the policy change, we make three sample restrictions. First, we include in our analysis sample only individuals who received an arrear for non-payment in 2000 or 2001, and thus had those non-payment 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 (before the policy change). Note that all individuals in our final analysis sample made their non-payment (and subsequent payment) decisions under the same beliefs about the Swedish credit registry data retention policies. Thus, the actions

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

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 pre-determined decisions. Third, because of the bi-monthly nature of the credit registry data (December-January, February-March, etc.), we restrict our sample to defaults occuring strictly after January 2000. The December-January bi-monthly flow of removals considers individuals whose information was deleted as of the previous year (close to exactly three years after it occured), 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 bi-month. 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. This 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 non-payments were reported for the individuals in our sample. In particular, Control group individuals whose non-payment 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 October 2003, while Treated group individuals whose non-payment were recorded in the first months of the year were reported in the credit registries for exactly three years.

We compare the probability that an individual from the Treated group is employed (as well as other outcomes such as income and self-employment) the year her non-payment is deleted from the credit registry to the same probability for a Control group individual. We control for default-year cohort effects by comparing early-in-the-year versus late-in-the-year defaulters for both groups.

We define the treatment groups with the variable $treated_i$, which equals one if borrower i's last non-payment occured during 2001 and zero if it occured during 2000. We interact $treated_i$ with the dummy variable $early_i$, which distinguishes between individuals whose non-payments occured early and late during the year. Because in our data each individual is assigned to a bi-monthly cohort of defaulters, we define $early_i$ to equal one for those individuals whose last non-payment occured in February-March or April-May, and zero for individuals whose last non-payment occured in August-September or October-November.[10] Finally, we create a dummy called $post_{i,t}$ which equals one the year borrower i's non-payment signal is removed (2003 for the Control group, 2004 for the Treated group). In that sense, the variable $post_{i,t}$ is measured in event time, which is defined starting at 0 in 2000 for the Control group and in 2001 for the Treated group. Thus, event time year 3 represents the year in which individual's non-payment flag is deleted from the credit bureau.We include individual fixed effects ω_i, year fixed effects ω_t, and event time fixed effects ω_τ. Our main specification is the following triple differences regression:

$$\begin{aligned} employed_{i,t} &= \omega_i + \omega_t + \omega_\tau + \beta treated_i \times early_i \times post_{i,t} + \delta post_{i,t} + \\ &\quad \gamma treated_i \times post_{i,t} + \lambda early_i \times post_{i,t} + \varepsilon_{i,t}. \end{aligned} \tag{3.1}$$

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 below, measures the differential probability of being employed for the Treated and Control group, for individuals whose non-payment was reported early in the year relative to those whose non-payment was reported late in the year, the year(s) after each individual's non-payment is no longer reported relative to the three prior years. The coefficients δ and λ capture differences in employment for individuals with late and early in the yeardefault, respectively, after their information no longer remains publicly available. Finally, γ captures differential employment trends for all Control group individuals after their information is no longer reported.

3.4. **Summary statistics.** Before presenting the regression output, we discuss 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. Table 3 presents information on our outcome variables. The top panel presents a brief definition of 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 6 years (2000 to 2005). Our summary stats are estimated the three years before these individuals' non-payment flags are removed, which correspond to 2000 to 2002 for the Control group and 2001 to 2003 for the Treated group.

[10] 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.

During those years, on average 43% percent of individuals in our sample are employed during the full year, while 79% received some positive wage income. We transform our income measures, which are in units of hundreds of SEK, to logarithms and add a 1 to include the effect of zero income. On average, log(income + 1), the log of pre-tax income, equals 5.6, which corresponds roughly to a pre-tax income of 102,000 SEK or $12,200. This figure is approximately 20 percent lower than the mean of the general population during that same period. 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 five percent of all individuals in our sample appear as self-employed. Further, 15% of individuals change address during a year (defined as changing kommun, a geographic division akin to an MSA–there are roughly 200 kommuns in Sweden). Finally, in terms of demographic characteristics, individuals in our sample are 42.8 years old, 60% male, and 6% own a house.

4. The Effect of Credit Information on Employment

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

4.1. **Graphical evidence.** The identification assumption for regression (3.1) 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 0 in 2000 for the Control group and in 2001 for the Treated group. There are no noticeable 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 non-payment 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. The figures also hint at our main results: Treated group individuals who defaulted earl in the year exhibit a higher probability of employment and earn higher incomes after their non-payment flags are removed.

4.2. **Main results.** Table 3 presents the coefficient of interest of specification (3.1). 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 fully employed during each year. Column 1 documents that the probability that an individual whose information is reported for a shorter period is employed increases by 2.8 percentage points , the year their non-payment is removed from the credit registry (year 3). This effect represents a 6.5% increase relative to the pre-period average employment rate (43%). 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 income from work during the year. Indeed, Column 4 shows that Treated group individuals are three 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.

We further exploit our empirical setup to explore the impact of credit market information on additional labor market outcomes. Columns 1-3 of Table 4 display the output of our main regression model (3.1) during the two years after removal of the non-payment 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 pre-tax income, log(income + 1). Income measures are in hundred of SEK. To capture the effects of zero income from employment in the logarithms, we replace zeros with a one.[11]

Consistent with our previous results, we find that individuals whose non-payment flag was retained for less time earn higher wage incomes. This effect 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

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

estimate that approximately 54% of the higher wage effect is driven by the extensive margin.[12]

We explore whether our measured effect on wage income varies for individuals who were differentially exposed to the policy change, measured by the difference in the number of months that Treated and Control group individuals' past defaults were publicly available. In the bottom panel of Figure 4 we repeat the exercise from regression (4.1), replacing wage income as the outcome variable. The Figure shows once again that, consistent with our identification assumption, Treated group individuals who were more exposed to the policy change (who defaulted early in the year) have monotonically higher incomes than Control group individuals, relative to those who were less exposed.

Column 2 of Table 4 show that individuals whose non-payment flag is publicly available for less time are one 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.[13] Indeed, note that individuals whose non-payment flag is retained for longer become self-employed in spite of the fact that they are relatively more credit constrained. This evidence suggests that in our setting, the credit information channel affects entrepreneurial activity above and beyond a story based on credit constraints.

Finally, Column 3 shows that credit market information affects borrower's total pre-tax income in a significant manner. That is, individuals' total incomes are higher when their information on past repayments is 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

[12]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% intensive margin effect from Column 4 in Table 3 corresponds to an 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 12 in the Internet Appendix). Thus, the extensive margin represents a $\frac{2{,}129}{3{,}987} = 53.4\%$ of the total wage effect. In unreported results, we find that the point estimate of the regression using logarithm of wages as outcome, without replacing zeros for ones, is positive but insignificant, which suggests a small effect on the intensive margin of wages. We interpret this result cautiously as such regression is ran on a selected sample of individuals who were employed before and after the policy change. Latent wages of unemployed individuals may bias this correlation.

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

fact that individuals are able to attenuate part of the effect of credit information on employment through self-employment.

4.3. **Results by treatment intensity.** Our identification strategy relies on variation in the retention times of non-payment 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 different labor market outcomes.

We proceed by categorizing individuals in our sample who were exposed to longer retention times in four groups by the bi-month in which their non-payment occured: 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 fewer number of months by which the non-payment flag was available in the credit bureaus, for Treated relative to Control group individuals: the August-September cohort has 1 month average exposure, June-July has 3 months average exposure, April-May has 5 months average exposure, and February-March has 7 months average exposure. If information about non-payments affects the probability of being employed, we hypothesize that the measure of months of exposure should be positively correlated with the probability of being employed during a given year. Note that the October-November cohort has, by construction, zero months exposure.

To test this hypothesis, we modify regression model (3.1) 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. In practice, this categorizes individuals by their bi-month of default. Thus, we test the following specification:

$$\begin{aligned} 1\left(wage>0\right)_{i,t} &= \omega_i+\omega_\tau+\omega_t+\sum_{t=1,3,5,7}\beta_t 1\left(exposuremonths_i=t\right)\times treated_i\times post_{i,t}+ \\ &\quad \delta\times post_{i,t}+\gamma treated_i\times post_{i,t}+ \\ &\quad \sum_{t=1,3,5,7}\lambda_t 1\left(exposuremonths_i=t\right)\times post_{i,t}+\varepsilon_{i,t}. \end{aligned} \tag{4.1}$$

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 $1\left(wage>0\right)_{i,t}$ as the outcome, and limit the post period to the year during which the non-payment flag is removed. Figure 4 shows a plot of the regression coefficients β_t and associated 95% confidence interval of this regression. Consistent with our identification assumption, the measured effect is stronger for individuals who were exposed to longer retention times due to the month in which their default occured. Further, the pattern is monotonic for 3, 5, and 7 months of exposure. The pattern is very similar for $log(wage)$, also shown in Figure 4. One month of exposure corresponds to a decrease of 0.14 log wage points, while 7 months of exposure corresponds to a decrease 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:

$$\begin{aligned} 1\left(wage>0\right)_{i,t} &= \omega_i+\omega_t+\omega_\tau+\beta exposuremonths_i\times treated_i\times post_{i,t}+ \\ &\quad \delta\times post_{i,t}+\gamma treated_i\times post_{i,t}+ \\ &\quad \sum_{t=1,3,5,7}\lambda_t 1\left(exposuremonths_i=t\right)\times post_{i,t}+\varepsilon_{i,t}. \end{aligned} \tag{4.2}$$

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 inccrease 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(wage + 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% level. We believe that the results in both Figure 4 and Table 5 are consistent with and provide credibility for the identification assumption.

4.4. **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 less able to move if negative information is held by the credit bureau for longer. 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's municipality has changed 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, on a baseline mean of 14.8%. While a large effect in 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%) 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 pre-period. 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% level, we find them highly suggestive of a type of mobility lock in the rental market due to credit market information.[14]

Second, we ask whether some individuals respond to decreased labor market opportunities by seeking additional schooling. When wage jobs become more scarce, the opportunity cost of schooling decreases, which may in turn increase the demand for schooling. 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 schooling is indeed one margin of 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% 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 leaving the labor force, turning to self-employment activities, and seeking additional schooling. 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.

4.5. **Incidence.** We document large positive impacts of decreased retention time of non-payment flags on wage employment and income for our sample of previous defaulters. It is natural to next ask, for which types of individuals are the effects strongest? First, we explore whether the effects differ by the employment history, namely the pre-period (event time 2) employment status. There is reason to believe

[14]This is similar to the housing lock-in documented by Struyven (2014) in the case of Dutch homeowners. Struyven (2014) argues that individuals with higher loan-to-value ratios on their mortgages and lower home equity experience reduced mobility.

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 Section 2. Second, the condition of being financially constrained that 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 on 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 searching actively 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 (i.e., Kroft, Lange, and Notowidigdo (2013)) and may have even stopped searching actively. 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 pre-period versus those who were employed. In Column 1 we run our main specification (Equation 3.1) 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 remarkable 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 pre-period years. The non-chronic unemployed are those who are unemployed at event time 2 who worked more than 1 year in the three pre-period 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 less years of completed schooling (the median number of years), 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.

One interpretation of the heterogeneity by years of schooling 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 is less relevant than others and as such it is down weighted in the employer's inference problem. On the other hand, individuals with little formal education have fewer ways to signal their type. As a result, their past defaults are a stronger signal of future productivity.

5. Interpretation

We have documented statistically and economically large effects of public credit information on employment and other outcomes. In Section 2 above we explain the two broad categories of mechanisms that could give rise to a relationship between credit market information and labor market outcomes: labor demand and labor labor supply. The former may correspond to employers using the signal of past default to screen employees. Such screening could be motivated by unobserved fixed differences in productivity correlated with credit information or by a causal effect of reduced access to credit on productivity. There is also a range of models that could result in links between credit information and labor supply, most of which are mediated by the credit supply channel observed in Bos and Nakamura (2014). Finally, negative credit information may limit geographic mobility if landlord screen tenants based on credit checks, which may also affect labor supply. Here we exploit the richness of our data to explore which of the possible mechanisms is most consistent with our empirical findings.

We start by addressing the plausbility of the different labor supply narratives. Several pieces of evidence suggest that this is not the main mechanism to reconcile our findings. First, we can rule out a large class of labor supply theories with our headline result. Recall that models with a precautionary labor supply motive predict a negative correlation between credit supply and labor supply. In contrast, our results indicate a positive relationship between the removal of negative credit information and employment.

Second, our results also appear at odds with a simple story where improved credit information allows individuals to invest in supplying their labor. Note that in Table 4 we document that individuals whose credit information is retained for fewer months are less likely to be self-employed. In particular, 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, it is hard to argue that individuals are constrained to pay a fixed cost to enter labor markets and are at the same time unconstrained to pay a fixed cost to become self-employed.

Further, we can measure credit record inquiries by non-financial institutions. These institutions include utilities companies, legal representation, and crucially, employers. 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 (3.1) using the number of non-financial credit checks as the outcome should be positive and significant. Column 1 in Table 9 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 non-payment is removed from the credit bureau. However, job search behavior does not appear to change.

Finally, 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 withat least 5,000 inhabitants (there are 290 kommuns in Sweden). In Table 10 we present our regression results, run in separate samples depending on whether the individual's pre-period kommun of residence level of unemployment is higher or lower than 3.85% (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.[15] Columns 4 through 6 show that wages follow the same pattern. However, columns 7, 8, and 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 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 and low unemployment areas, but only in the latter does this have a differential effect on the probability of being employed. Consequently, the effects of credit market information on labor market outcomes in our setting are likely not explained by changes in an individual's labor supply, but rather by changes in the demand for these individuals' labor conditional on their credit market information.

We also view the results in Table 10 as being consistent with a labor demand story. 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.

Finally, the results on mobility in Table FIXME are unable to explain the full effect.Although the lack of mobility to better labor markets induced by bad credit information may also have a causal role explaining our results, quantitatively it may explain at most 27% of the baseline effect of information on employment in Table 3. [16]

We conclude that in our setting, the causal effect of the removal of information on past defaults in employment is driven by its effect on labor demand, and specifically

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

[16]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.

on the fact that employers screen individuals based on their past defaults as a signal of unobserved heterogeneity in productivity. Within the theories of how information may affect the demand for labor, we note that the variation of our results by local labor market tightness is better explained by a story where past defaults is used as a signal for unobserved persistent heterogeneity in productivity. This result is consistent with Kroft, Lange, and Notowidigdo (2013), who find that unemployment duration has a stronger causal effect on future employment in tight labor markets. Indeed, the causal or temporary effect of access to credit on an individual's productivity is not likely to vary by local labor market tightness. Further, the value of the signal is likely to be higher when labor markets are tight, i.e., when the average quality of the pool of unemployed individuals is lower.

6. Conclusion

We combine a unique natural experiment in Sweden with very 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 her success in the labor market. In particular, we find robust evidence that negative credit information makes individuals less likely to be employed, and as a result, they earn lower incomes. The large incidence of credit information errors and of negative marks from various sources such as medical debts adds to the potential damaging effect of the use of information in hiring decisions.[17] We also show that when individuals are unemployable due to their negative information, they adjust in part by turning to self-employment as a seemingly-inferior alternative source of income.

Our results suggest that the consequences of default through information sharing in the credit bureau have profound effects on the livelihoods of individuals. While ex ante, such effects strengthen the incentive mechanism of the credit bureau, a temporary shock causing an individual to default may have lasting and profound consequences. During default episodes (i.e., financial crises), such spill-over effects may, under certain conditions, serve to amplify and exacerbate downturns. Extrapolating our results to a different setting and market, our results may help explain why so few individuals defaulted in their mortgages when their equity values became negative during the housing crisis (Foote, Gerardi, and Willen (2008)).

[17]E.g. 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/. We acknowledge that medical debts are be more relevant for the US context than the Swedish context, given Sweden's universal health care programs.

References

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

Attorney General, State of New York, 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*.

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*.

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.

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.

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

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, 2008, 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*.

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.

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.

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

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

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.

Myers, Stewart C, 1977, Determinants of corporate borrowing, *Journal of financial economics* 5, 147–175.

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

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 non-payment flag for individuals with non-payments in different moments of the year. In particular, the figure shows that individuals whose non-payment occured early 2001 had a reduced retention time of past non-payments. In contrast, individuals whose non-payment occured early in 2000 were reported in the credit registries until October 2003.

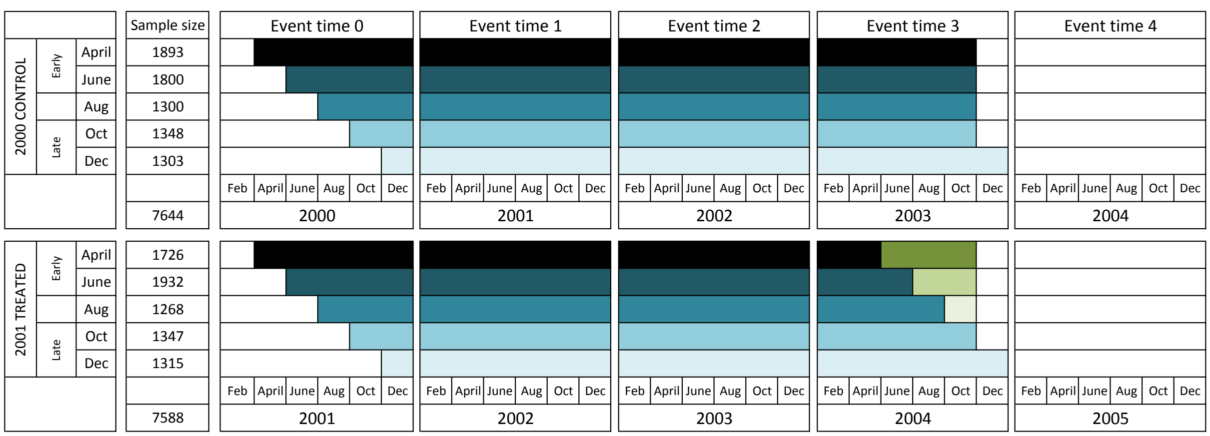


Figure 2. Frequency of removal of non-payment flag over time
This figure displays the distribution of the removal of non-payments over time. In the old regime the credit bureau removed all negative arrears that were eligible for removal once a year, on December 31. Each non-payment was eligible for removal in the third year after the year in which it was received. Because of the bi-monthly feature of our data, and because removals are inferred as differences in the stock of reported defaults, these non-payments corresponds to the February-March bi-month (labelled February). This regime ended in October 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.

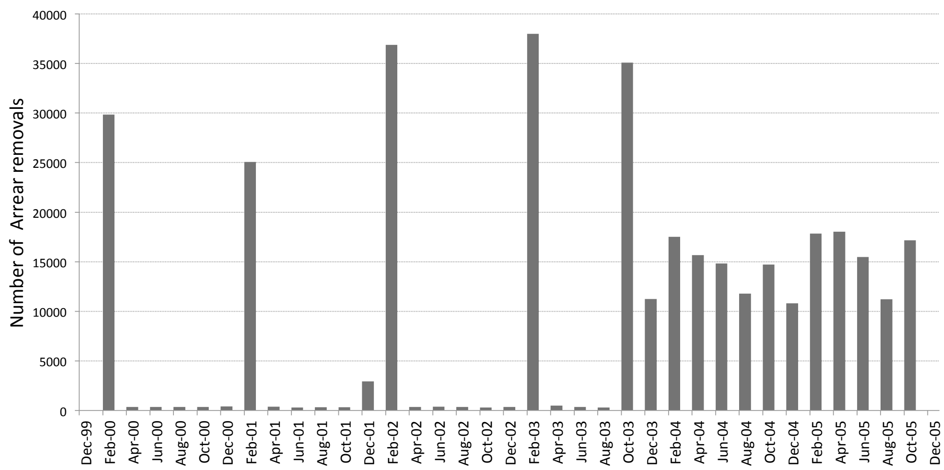


FIGURE 3. Pre-trends

This figure shows that there is no difference in the pre-period trends (before the policy change) of the difference between Early and Late defaulters, in the Trated and Control group for our main outcomes. The top panel shows pre-period trends for *employed* and 1(wage>0), which equals one if an individual received any wage income, the lower panel for log(wage+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.

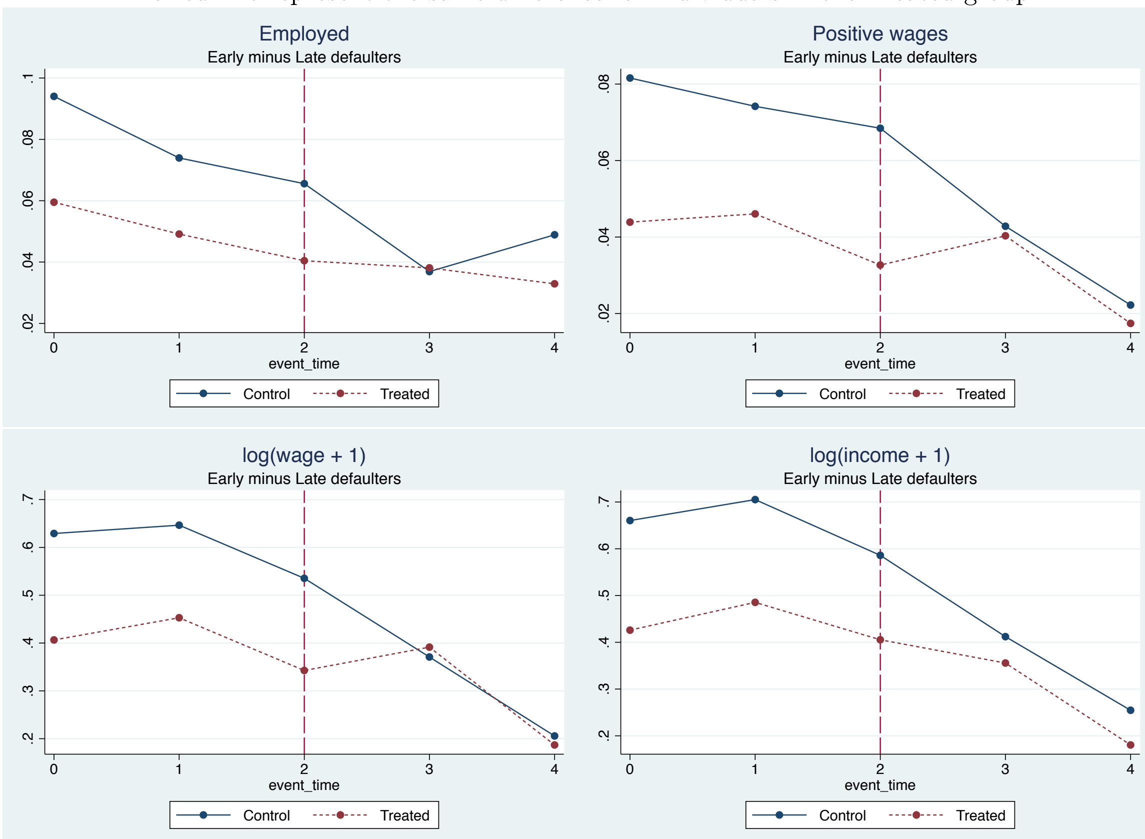
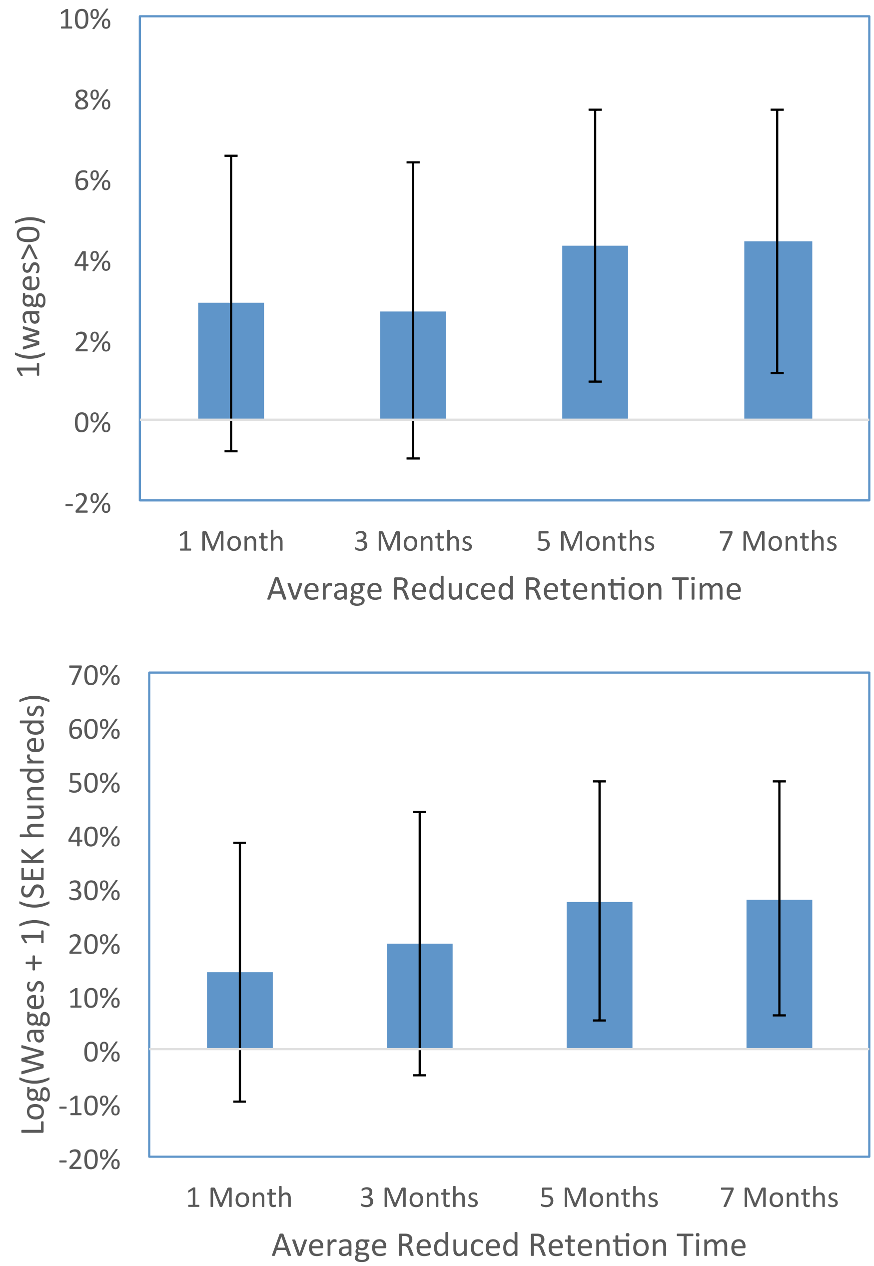


Figure 4. Retention time exposure and employment status

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 (4.1)) 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:

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

Employment is defined as a positive wage income, documented by tax records, $1\left(wage_{i,t} > 0\right)$. 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 non-payment flags. Data is a yearly panel of a random sample of the universe of Swedish individuals with a credit score, between 2000 and 2012. Standard errors are clustered at the individual level. *, **, and *** represent 10, 5, and 1 percent significance level, respectively.

	(1)	(2)	(3)	(4)
Dependent variable	$1\left(wage_{i,t} > 0\right)$	$1\left(wage_{i,t} > 0\right)$	$1\left(wage_{i,t} > 0\right)$	$1\left(wage_{i,t} > 0\right)$
$creditscore_{t-1}$	-0.0559***		-0.0191***	
	(0.0018)		(0.0012)	
$creditscore_{t-2}$		-0.0524***		-0.0164***
		(0.0018)		(0.0011)
Controls	YES	YES		
Ind FE			YES	YES
Obs	515,083	499,361	515,083	499,406
R^2	0.0732	0.0689	0.0486	0.0469
Individuals	15,682	15,682	15,683	15,683

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 year before non-payments are deleted, which corresponds to 2002 for the cohort that defaulted in 2000 (old regime), and 2003 for the cohort that defaulted in 2004 (new regime).

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 pre-tax income, in 100 SEK; zeros replaced by 1.
log(wage + 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.

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(wage + 1)	5.57	2.97	7.04
Self-employed	0.05	0.21	
Relocates	0.15	0.35	
Years of schooling	10.70	1.76	11
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	$1\,(wage > 0)$	$1\,(wage > 0)$	$1\,(wage > 0)$
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)	(4)
Coefficient	log(wage + 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 non-payment flags on the probability of receiving any wage income during the year. The table shows contains the coefficient β from regression:

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

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	$1\left(wage > 0\right)$	$1\left(wage > 0\right)$	log(wage + 1)	log(wage + 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 pre-period employment history

This table shows differential effects of credit information on employment depending on pre-period employment status. The table shows the regression output of our main regression model (3.1) 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 non-payments 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 pre-period. 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	$1\,(wage > 0)$	$1\,(wage > 0)$	$1\,(wage > 0)$	$1\,(wage > 0)$
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(wage + 1)	log(wage + 1)	log(wage + 1)	log(wage + 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 9. Treatment Effects on Number of Credit Inquiries: Financial and Non-Financial

The table shows the regression output of our main regression model (3.1) for different sub-samples. 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.

	(1)	(2)
Coefficient	non-financial inquiries	financial inquiries
β	0.0035	0.1256**
	(0.030)	(0.057)
Pre-period 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

Table 8. Heterogeneity by pre-period education levels

This table shows differential effects of credit information on employment depending on pre-period level of education. The table shows the regression output of our main regression model (3.1) 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(wage+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	$1\,(wage > 0)$	$1\,(wage > 0)$	$log(wage + 1)$	$log(wage + 1)$
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 10. 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 (3.1) for different sub-samples. Columns 1restricts 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	$1\,(wage > 0)$	$1\,(wage > 0)$	$1\,(wage > 0)$
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(wage + 1)	log(wage + 1)	log(wage + 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

Supplemental Appendix

Table 11. 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 12. 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 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	wage
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