Leaving Your Past Behind: Debt Relief and Examiner Instrumental Variables*

Jakob Beuschlein†

August 14, 2024

Abstract

Many countries provide debt relief programs for over-indebted individuals. I use an examiner leniency design to estimate the effects of the Swedish debt restructuring program on subsequent labor market outcomes. I find that participation in the program has, on average, negative effects on earnings and employment. These findings mask large heterogeneity. While accepted high-income applicants experience increases in income and employment, low-income households face substantial negative effects that persist even after the program ends. The combination of high adjustment rates in the repayment plans for participants and wage garnishment for rejected applicants can explain these disparities. I calibrate a labor supply model and show that modest changes in the structure of the repayment plan can improve welfare for both debtors and creditors. Another contribution of this paper is to show that if there is path dependency in examiners' decision-making, classical examiner instrumental variables (IV) based on leave-one-out estimates of examiner leniency will be biased. I report substantial biases relative to the OLS bias in the setting of the Swedish debt restructuring program. Instead, I propose a past-cases-only estimator that is robust to path-dependent decision-making.

JEL Classification: C26, D14, J22, J64, K35

Keywords: Debt relief, personal bankruptcy, unemployment, examiner instrumental variables

*I thank Ulrika Ahrsjö, Monir Bounadi, Mitch Downey, Peter Hull, Markus Jäntti, Bouke Klein Teeselink, Juan Llavador Peralt, Patrizia Massner, Thomas Mikaelsen, Jens Oehlen, David Schönholzer, David Seim, Jósef Sigurdsson, David Strömberg, Jens Wikström, Quan Cheng Xie, Pablo Zárate, and conference and seminar participants in Brussels, Stockholm and Uppsala for helpful comments.

†Department of Economics, Stockholm University. Email: jakob.beuschlein@su.se

1 Introduction

Debt relief programs worldwide aim to offer individuals overburdened by debt a chance to reset financially and to improve their overall life circumstances. Starting in the 1990s and accelerated by the 2008 financial crisis, many countries adopted or significantly adjusted their personal insolvency or debt relief regimes.[1] The objective of these reforms was to create legal frameworks for the orderly resolution of debts, while balancing social policy considerations with maintaining credit discipline (Liu and Rosenberg, 2013). Although there has been an international trend towards partial repayment plans that also allow so-called 'No Income No Asset' debtors to discharge their debt (Ramsay, 2017a), the strictness of these programs varies considerably across countries, both in terms of acceptance criteria and the conditions under which debts can be discharged. The Swedish debt restructuring program is, unlike the country's generous welfare system, considered very strict by international comparison. Accepted applicants are economically worse off than their peers in other countries and debts are only discharged after successfully completing a stringent five-year repayment plan (Ramsay, 2017b).

I use the quasi-random assignment of case examiners to debt restructuring applications at the Swedish Enforcement Authority (SEA) and a novel examiner instrumental variable (IV) design, which estimates leniency based on past cases only, to estimate treatment effects of this program. While previous work in the US and Denmark finds that participants are substantially better off economically after gaining access to these programs (Dobbie and Song, 2015; Bruze, Hilsløv, and Maibom, 2024), I find negative effects on labor income and employment in the seven years after the program for accepted applicants in Sweden. Labor income falls by 9.1 percent and employment decreases by 14 percent, relative to rejected applicants.

These averages, however, mask large heterogeneity, most pronounced along initial income of applicants. High-income participants experience increases in income and employment of 16.8 percent and 4.1 percent. In contrast, applicants who are initially worse off are 64.2 percent less likely to work and earn 61.9 percent less. These disparities persist two years after the program has ended. Additionally, I provide further evidence that these differences in income and employment levels remain up to nine years after the program's start for both high and low-income participants by estimating matched difference-in-differences (DiD) event studies.

I then estimate treatment effects for the marginal applicant at each level of examiner leniency. The marginal treatment effect framework (Heckman and Vytlacil, 2005) allows estimating treatment effects against an individual's unobserved resistance to treatment. This unobserved resistance can be mapped to the leniency of the examiner that decides the current application. Treatment effects on income and employment are negative for the most lenient examiners and positive for the strictest examiners. This finding has implications for the overall effects of ex-

[1] Liu and Rosenberg (2013) and Bergthaler, Garrido, and Rosha (2023) provide details on such reforms in several European countries after the financial crisis. Russia introduced its personal bankruptcy law in 2015, while India significantly changed its laws in 2016. China followed by introducing a personal bankruptcy regulation in the Shenzhen Special Economic Zone in 2020. In the US, where consumer debt relief has a longer history, almost 10 percent of households have filed for personal bankruptcy at some point (Stavins, 2000).

panding the debt restructuring program. An expansion that allows more applicants to discharge their debt effectively increases leniency among certain examiners. The overall impact will depend on where in the leniency distribution changes occur. Initially strict examiners will accept more applicants with negative marginal treatment effects, while already lenient examiners will accept applicants with positive treatment effects. In summary, participation in the debt restructuring program appears to hurt a substantial share of applicants, especially those who are already economically worse off.

The debt restructuring program requires participants to repay parts of their debt to their creditors depending on their financial capacities over a period of typically five years. The program's repayment plan is set at the outset and allows participants to live at a subsistence level. This structure implies that a high share of low-income debtors can discharge their debts without any repayments under the initial repayment plan. If participants' financial circumstances change permanently due to unforeseeable events, such as large increases in labor income after transitions from non-employment to employment, repayment plans can be adjusted if creditors apply for adjustments. These adjustments allow participants to keep 4,000-5,000 Swedish krona (SEK) per month.[2] Any additional income exceeding this threshold will be added to the repayment plan and thus effectively taxed at a 100 percent rate.

Delinquent debtors who do not participate in the debt restructuring program can be subject to a wage garnishment regime in which their employers transfer parts of their salary to the SEA which in turn uses these funds to pay back their creditors.

The interplay of wage garnishment and potentially high adjustment rates in the repayment plan schedule can explain the divergence in results for high and low-income participants by creating adverse labor supply incentives for both *rejected* high-income applicants and *accepted* low-income applicants. High-income participants work on average 35 hours per week the year before applying to the program. This allows them to benefit from normal wage increases without risking any adjustments to their repayment obligations. Rejected high-income applicants, on the other hand, face relatively high wage garnishment rates reducing work incentives if substitution effects outweigh income effects. Low-income applicants, in contrast, are substantially less likely to work, with more than 90 percent of them having no employment in the year before applying. Seeking employment will, therefore, almost always lead to large discrete jumps in income which can result in severe adjustments to the repayment plan. Although rejected low-income applicants can also face wage garnishment, only a relatively small share does empirically and their average garnishment rates are low. This creates large incentives for low-income participants to avoid seeking work during the program. After the program ends, several mechanisms can further explain the persistence of the effects. Human capital depreciation (Dinerstein, Rigissa, and Yannelis, 2022) and psychological detachment can reduce incentives to seek employment. Furthermore, employers might discriminate against job seekers with long periods of

[2]Throughout the paper, I will report results in 2019 prices. The average exchange rate to the US dollar in 2019 was 9.46 SEK.

non-employment (Kroft, Lange, and Notowidigdo, 2013; Eriksson and Rooth, 2014).

I calibrate a dynamic model of labor supply under debt restructuring and show that these mechanisms can explain the empirical results. Initially employed applicants will increase their labor supply relative to rejected applicants who face wage garnishment. After exogenously setting initial employment to zero for accepted applicants, the high adjustment rates will prevent re-employment. Human capital depreciation prevents re-entry into the labor force after the program ends after five years. I then use the model to simulate policy reforms by reducing the adjustment rate that participants face. Already a decrease from 100 to 80 percent is sufficient to induce initially unemployed participants to seek part-time employment, which further translates into higher employment after the program ends. The additional benefit over seven years amounts to 1.4 million SEK or 5.1 times the respective annual full-time income. This additional benefit is distributed among higher consumption for debtors, increased government revenues from reduced social benefit payouts and higher income taxes, and greater repayments to creditors, with the first two outweighing the latter in quantitative terms. I further show that lower adjustment rates are unlikely to create additional labor supply moral hazard among applicants who might otherwise be incentivized to reduce their initial labor supply to lessen repayment obligations.

My findings highlight potential pitfalls in designing debt restructuring programs that place relatively high weight on concerns over debtor moral hazard or creditor welfare. This is especially relevant, as Indarte (2023) finds that increases in the probability of debt forgiveness in the US have only small effects on debtor moral hazard.

Another contribution of this paper is to show that examiner or judge IV designs[3] rely on the assumption that there is no path dependency in examiner decision-making such that current decisions do not depend on past case characteristics. This assumption is, however, at odds with a large literature in behavioral economics and psychology showing evidence for path dependencies in sequential decision-making across many different settings (e.g. Chen, Moskowitz, and Shue, 2016; Jin, Tang, Ye, Yi, and Zhong, 2023). I show that classical examiner IV estimations relying on jackknife or leave-one-out estimates of leniency can be inconsistent with path-dependent decision-making.[4] Intuitively, constructing the instrument from future cases which in turn can depend on current case characteristics will re-introduce endogeneity into the estimation. Instead, I propose to construct the leniency instrument from *past cases only*. I evaluate the bias of the leave-one-out instrument in simulations and show that this can impact estimates in the empirical setting of debt restructuring. Specifically, the treatment effect on income rises from -17 percent to -9.6 percent whereas the treatment effect on employment increases from -15.6 percentage points to -9.9 percentage points. These changes correspond

[3] Kling (2006) pioneered this approach, which has since gained popularity in applied microeconomics. For recent applications see for example Dahl, Kostøl, and Mogstad (2014) on intergenerational welfare cultures, Gross and Baron (2022) on foster care, Galasso and Schankerman (2015) on cumulative innovation, Grindaker, Kostøl, and Merkle (2024) on firm bankruptcy, Dobbie, Goldin, and Yang (2018) on pre-trial incarceration, or Collinson, Humphries, Mader, Reed, Tannenbaum, and van Dijk (2024) on evictions and homelessness among others.

[4] This issue is distinct from problems arising through clustered sampling, as in Chao, Swanson, and Woutersen (2023) and Frandesen, Leslie, and Mcintryie (2023)

to 45 percent and 91 percent of the OLS bias, respectively. I estimate average treatment effects (ATE) and demonstrate that the difference between the two instruments is, if anything, even larger when replacing 2SLS weighting schemes with constant weights. I provide further evidence for the presence of distorting path dependencies by showing that the leave-one-out instrument correlates with pre-defined applicant characteristics that predict treatment, while the past-cases-only instrument does not. I further show that using an instrument constructed from future cases yields estimates that are closer to or even exceed the OLS estimate compared to both the leave-one-out and the past-cases-only IV estimates highlighting the endogeneity of future cases.

The remainder of this paper is structured as follows. I describe the institutional context of debt restructuring in Section 2. In Section 3, I derive the asymptotic bias in examiner IV designs under path-dependency and develop a robust past-cases-only leniency estimator. I describe the data used and how I construct my estimation sample and instruments in Section 4. In Section 5, I report my main findings. In Section 6, I discuss evidence for the proposed mechanisms and simulate policy reforms using the calibrated labor supply model. Section 7 concludes.

2 Debt restructuring in Sweden

The debt restructuring system in Sweden was initiated in 1994 by the Debt Adjustment Act (*Skuldsaneringslagen*) with the objective of allowing participants economic rehabilitation (Lennander, 1991).[5] The relevant legal framework for this context is outlined in the 2006 version of the Debt Adjustment Act[6]. Overindebted applicants seeking debt restructuring can submit their application to the Swedish Enforcement Authority (SEA, *Kronofogden*) where an examiner (*Handläggare*) will decide on each case. Applications are free of charge and can easily be submitted online. The SEA is a government agency—originally established as a tax collection authority within the Ministry of Finance—that is responsible for debt collection, distraint, and evictions, among other things. Examiners within the SEA are normally, but not always, legal professionals with a university degree who have substantial discretion in decision-making (Larsson and Jacobsson, 2013).

Applications are evaluated based on two main criteria. First, applicants are not expected to be able to repay their debt in foreseeable future and secondly, debt restructuring is reasonable with regard to the debtor's personal and financial circumstances. Debtors who have difficulties fulfilling their obligations are often subject to a wage garnishment scheme administered by the SEA. Once accepted, any wage garnishment will be replaced by an individual repayment plan that typically lasts five years.[7] During these years debtors are required to repay parts of their debt to their creditors. The extent of repayment depends on the debtor's repayment capacity

[5]For a more detailed overview of the history of debt restructuring in Sweden see Ramsay (2017a).
[6]Skuldsaneringslag (2006:548).
[7]Shorter repayment plans may be accepted for older or sick applicants.

and a reserve amount that ensures they can maintain a subsistence level, considering personal circumstances such as the number of underage children in the household.[8] The repayment plan can be adjusted if the financial circumstances of the participant change permanently due to unforeseeable events, such as obtaining higher salaries from new employment. Normal wage increases do not qualify as such events. Generally, these adjustments can occur if the debtor's monthly income decreases by more than 500 Swedish Krona (SEK) throughout at least three months or if it increases by more than 4,000 to 5,000 SEK. To initiate changes, either debtors or creditors must apply to the SEA, where the application will be reviewed. The case worker responsible for this review will typically differ from the examiner who made the initial decision. The SEA does not inform creditors about improvements in debtors' financial circumstances. However, due to Sweden's extensive transparency laws, most relevant information is publicly available to creditors. For instance, information on debtors' incomes can be accessed through the Swedish Tax Agency. Typically, the debtor is allowed to keep any increases in income up to 4,000-5,000 SEK. Any additional income above this threshold will be included into the initial repayment plan. Adjustments of the repayment plan can be severe if the increase in the debtor's income is large. Consider, for example, a working-age debtor living alone who is unemployed in 2015, but secures a job paying at least 10,000 SEK per month in 2016. The average monthly disposable income for such individuals is 6,932 SEK when unemployed and rises to 17,160 SEK once employed.[9] This exceeds the reserve amount that the SEA considers as a subsistence minimum.[10] Debtors are allowed to keep 4,000-5,000 SEK of the increase of 10,228 SEK which corresponds to an implicit tax rate between 51.1 percent and 60.9 percent.[11] After the repayment plan is fulfilled the debtors are forgiven their unsecured debt except for obligations due to family law.

The assignment mechanism for applications to examiners is based on a national register of all cases. Examiners with free capacities will be assigned the oldest of the remaining cases. This quasi-random allocation of case to examiners allows me to use variation in examiners' leniency to provide causal estimates. There are two steps in the decision-making process. In a first, more bureaucratic, decision applications who do not fulfill basic criteria, for example if applicants' debt burden is too low given their age, will be rejected.[12] Remaining applicants will then be asked to hand in documentation after which their case will be thoroughly reviewed by the examiner who then makes a final decision. I will use data on these final decisions. Note that

[8] Around 40 percent of participants do not have to repay anything according to their initial repayment plan. The remaining 60 percent of applicants must repay 26 percent (174,000 SEK) of their debt on average.

[9] The numbers are derived from the Longitudinal Integrated Database for Health Insurance and Labor Market Studies covering the entire Swedish population.

[10] The basic reserve amount is 6,090 SEK for single households, but may be higher depending on costs for accommodation. Reserve amounts can easily be computed via an online calculator at the SEA's website.

[11] This already takes into account taxes and changes in social benefits. The corresponding average gross labor income is 17,748 SEK.

[12] Applicants can self-assess whether they meet the requirements for receiving debt settlement through an online test. For example, a 45-year-old single applicant earning 25,000 SEK per month, who did not accumulate debt through gambling, has a good chance of being considered for debt restructuring with a debt burden of 750,000 SEK, but not with a debt burden of 250,000 SEK.

if the first decision was also based on the examiner's leniency, this could induce a correlation between case characteristics and my instrument. This is, however, unlikely to be an issue given the standardized nature of the initial decision. In section 4.3.1 I show that the instrument is uncorrelated with observable applicant characteristics such as age or income that are highly predictive of being accepted to the program. If initial decisions were a function of examiner leniency we should find imbalance in the instrument.[13] Another threat to identification would be if examiners directly influence debtors' outcomes and this influence is systematically related to their leniency. While fundamentally untestable I do not expect this to be an issue in the given setting for two reasons. First, the SEA is supposed to function as a 'neutral intermediary' between debtors and creditors and should not intervene to favor either side. For this reason, the SEA, for example, does not provide advice to debtors about the validity of claims against them (Ramsay, 2017a). Secondly, while successful applicants do receive advise on how to improve their situation as part of the debt restructuring program, this is administered by either the respective municipality or the Swedish Consumer Agency (*Konsumverket*) and not the SEA.

3 Identification - examiner IV and path dependency

3.1 Path dependency in sequential decision-making

The general idea behind examiner leniency designs is the belief that decision makers will inherently differ in their propensity to make decisions of a certain kind. This is often summarized by a hypothetical exogenous leniency parameter denoted as L_j, where j refers to the decision maker.[14] A simple decision model consistent with classical examiner IV can be described as follows

$$D_{jt} = d^c\left(L_j, U_{jt}\right). \tag{1}$$

Here, t is a chronological index for the specific case, and U_{jt} summarizes all relevant case-specific information upon which the examiner bases their decision. Importantly, the examiner considers only contemporaneous case characteristics.

This model, however, is at odds with a large and growing literature in behavioral economics and psychology suggesting that, in reality, there may be path dependencies in human sequential decision-making. The literature proposes several psychological mechanisms for this phenomenon. For example, assimilation bias refers to the tendency of individuals to interpret new information in a manner that is consistent with previous beliefs. Bindler and Hjalmarsson (2019) report evidence for path dependencies in jury decision-making in English high-stakes criminal court cases in the 18th and 19th century that is consistent with assimilation bias. En-

[13] If leniency were correlated with case characteristics such that more lenient examiners deal with applicants who are relatively better off, we would expect the IV estimates for income and employment to be upward biased relative to the true effect.

[14] For simplicity, I will assume time constant leniency. All results below hold when allowing for time-varying leniency as long as the sources of variation are exogenous and there is some predictive power to induce correlation between different decisions of the same examiner.

glich, Mussweiler, and Strack (2006) show that legal experts' decisions in a lab experiment were similarly influenced by random anchoring. Srinivasan (2023) finds that judges in Cook County hand down sentences that are 97 days longer in the 10 days after a murder sentencing. Contrast effects, on the other hand, can lead to negative autocorrelation. If a particular case is perceived as exceptionally heinous, the subsequent case may appear less severe by comparison. Chen et al. (2016) find a negative autocorrelation in decisions of judges granting or denying refugees asylum. Bhuller and Henrik (2024) provide a different channel potentially inducing path dependency. They show that Norwegian judges presiding over criminal justice cases update their decision-making if previous decisions are reversed in appeal courts. If the probability of a decision being reversed depends on the characteristics of the respective cases, such feedback-based learning can create path dependencies.

The evidence on path dependency in decision-making is not only restricted to the judicial system. Other contexts in which decision may depend on previous alternatives are for example physicians' treatment decisions (Jin et al., 2023), job interviews (Radbruch and Schiprowski, 2023), TV talent shows (Page and Page, 2010), the judgement of attractiveness (Kramer, Jones, and Sharma, 2013), and refereeing in sports (Dohmen and Sauermann, 2016).

Following this literature, I propose a more general examiner decision model that also depends on the history of past case characteristics $\mathbf{U}_j^{k<t}$,

$$D_{jt} = d^p\left(L_j, U_{jt}, \mathbf{U}_j^{k<t}\right). \tag{2}$$

In the following section I will show that classical examiner IV can be inconsistent under such decision models.

3.2 Asymptotic bias in examiner IV

Let there be a set $\mathcal{J} = [1, \ldots, J]$ of examiners indexed by j[15] with cases $t \in \mathcal{T} = [1, \ldots, T]$ ordered chronologically. The structural model is

$$Y_{jt} = \beta D_{jt} + \alpha U_{jt} + \nu_{jt}. \tag{3}$$

Here, β denotes the treatment effect of an examiner's decision on an outcome of interest Y_{jt}. ν_{jt} is an iid structural error capturing causes of Y_{jt} unrelated to examiners' decisions. To simplify notation, I assume that there are no additional relevant covariates, including an intercept. In Appendix B1 I discuss extensions to models with covariates and heterogeneous treatment effects.

I treat Y_{jt}, D_{jt}, U_{jt}, and ν_{jt} as random variables while keeping J and T fixed. Let

$$\mathcal{P}_t = [t+1, t+P_t] \tag{4}$$

[15] Many empirical applications estimate leniency by examiner and year. In these settings, j refers to such an examiner-year combination.

be the set of future periods during which case characteristics U_{jt} potentially have an influence on decision D_{jk}. $P_t > 1$ and $P_t < T - t - 1$ depends on the underlying examiner decision-making model.

I assume that

$$E[D_{jt}\nu_{jk}] = 0 \text{ for all } k \in \mathcal{T}, \tag{A.1}$$

$$E[U_{jt}D_{jk}] = 0 \text{ for } k \notin \mathcal{P}_t \text{ and } k \neq t, \tag{A.2}$$

Assumption (A.1) asserts that the structural error is exogenous. Assumption (A.2) states that the current case characteristics are uncorrelated with either past decisions made without knowledge about case t or decisions sufficiently far in the future. Following the examiner decision model (2), I allow for correlation between the current case characteristics U_{jt} and the contemporaneous decision as well as decisions within the next P_t periods such that generally

$$E[U_{jt}D_{jk}] \neq 0 \text{ for } k \in \mathcal{P}_t \text{ or } k = t. \tag{5}$$

Let $\varepsilon_{jt} = \alpha U_{jt} + \nu_{jt}$ be the regression error. Since case characteristics U_{jt} are unobservable to the researcher, the endogeneity of current decisions to current and past cases (5) implies that a simple OLS regression of the outcome Y_{jt} on decision D_{jt} will not recover β. The corresponding OLS estimator is given by

$$\hat{\beta}_{OLS} = \frac{\sum_j \sum_t D_{jt}Y_{jt}}{\sum_j \sum_t D_{jt}^2}. \tag{6}$$

Throughout this section, I follow the literature on many instruments (Bekker, 1994) and evaluate the asymptotic biases by letting the number of examiners J grow to infinity while keeping the number of cases per examiner T fixed. Then by the weak law of large numbers as $J \to \infty$

$$\hat{\beta}_{OLS} \xrightarrow{p} \beta + \alpha \frac{\sum_t E\left[D_{jt}U_{jt}\right]}{\sum_t E\left[D_{jt}^2\right]}. \tag{7}$$

Given (5) the second term will be different from zero, because the examiner bases her decision on current case characteristics. To overcome this issue, many researchers use the leave-one-out average decision of examiner j as an instrument for the decision D_{jt}, to exclude the influence of the current case from the instrument.

$$Z_{jt}^{\text{LOO}} = \frac{1}{T-1} \sum_{k \neq t} D_{jk}. \tag{8}$$

Under random assignment of examiners to cases, and in the absence of path dependency in the decision-making process, this instrument will fulfill the independence assumption. To ensure the instrument's relevance, other cases of the same examiner have to have predictive power for the current decision

$$E[D_{jt}D_{jk}] \neq 0 \text{ for all } k \in \mathcal{T}. \tag{A.3}$$

Assumption (A.3) is sufficient, but stronger than what is strictly necessary in this setting. It allows to construct relevant instruments from any selection of other decisions an examiner makes and follows from decision models with time constant leniency such as (2).[16] The corresponding 2SLS estimator is given by

$$\hat{\beta}_{\text{LOO}} = \frac{\sum_j \sum_t Z_{jt}^{\text{LOO}} Y_{jt}}{\sum_j \sum_t Z_{jt}^{\text{LOO}} D_{jt}}. \tag{9}$$

This estimator is numerically equivalent to the jackknife IV estimator when using examiner fixed effects as instruments (Angrist, Imbens, and Kruger, 1999). As above when $J \to \infty$

$$\hat{\beta}_{\text{LOO}} \xrightarrow{p} \beta + \alpha \frac{\sum_t \sum_{k \in \mathcal{P}_t} E\left[D_{jk} U_{jt}\right]}{\sum_t \sum_{k \neq t} E\left[D_{jk} D_{jt}\right]}. \tag{10}$$

See Appendix B1.3 for technical details. Given path dependency (5) the second term will be different from zero since future decisions are a function of current case characteristics. The leave-one-out estimator will then be asymptotically biased.

Since the source of bias is the inclusion of future cases $t \in \mathcal{P}_t$, we can avoid this issue by constructing an alternative instrument from past observations only[17]

$$Z_{jt}^{\text{past}} = \frac{1}{t-1} \sum_{k<t} D_{jk}. \tag{11}$$

The respective 2SLS estimator is given by

$$\hat{\beta}_{\text{past}} = \frac{\sum_j \sum_t Z_{jt}^{\text{past}} Y_{jt}}{\sum_j \sum_t Z_{jt}^{\text{past}} D_{jt}} = \frac{\sum_j \sum_t \frac{1}{t-1} \sum_{k<t} D_{jk} Y_{jt}}{\sum_j \sum_t \frac{1}{t-1} \sum_{k<t} D_{jk} D_{jt}}. \tag{12}$$

Letting J grow to infinity we get that

$$\hat{\beta}_{\text{past}} \xrightarrow{p} \beta + \alpha \frac{\sum_t \sum_{k<t} E\left[D_{jk} U_{jt}\right]}{\sum_t \sum_{k<t} E\left[D_{jk} D_{jt}\right]} = \beta \tag{13}$$

where the last step follows from assumption (A.2). The past-cases-only estimator therefore recovers the true causal effect even under path-dependent examiner decision-making.

3.2.1 Monte Carlo simulations

In Appendix B2 I provide simulation results for a hypothetical examiner decision model which takes six past cases into account. The bias in the leave-one-out estimation increases if the importance of past cases for the current decision increases. For a fixed number of relevant past cases, the bias decreases in the number of cases per examiner. Perhaps surprisingly at first

[16] Since I abstract from treatment effect heterogeneity, I do not need to impose a monotonicity assumption here. See also Appendix B1.2. In the empirical part below, I will provide evidence that monotonicity holds in the context of debt restructuring.

[17] This instrument will only be defined for observations for which $t > 1$.

sight, the bias grows if the relevance of examiner leniency L_j for the current decision decreases or the relevance of current case characteristics U_{jt} increases. Both of these changes weaken the first stage which inversely scales the bias in (27).

3.2.2 Bias with many controls

In many empirical applications the leniency instrument is only valid conditional on controls such as court-by-year fixed effects. Kolesár (2013) shows that standard leave-one-out estimators can be biased in settings with many controls. He proposes a leave-one-out residualized unbiased jackknife IV estimator (UJIVE). In Appendix B1.1, I show that this issue is exacerbated under path-dependent decision-making. Intuitively, residualizing the instrument using observations whose application decisions depend on the current case characteristics can re-introduce endogeneity. A simple solution is to use leave-examiner-out residualization.

3.2.3 Heterogeneity in treatment effects

The general interpretation of the results derived above does not depend on the assumption of homogeneous treatment effects. However, when we allow for heterogeneity in treatment effects 2SLS estimates will identify a weighted average of treatment effects. Differences between the classical leave-one-out and the past-cases-only estimator can then arise due to biases from path dependency or differences in weighting schemes induced by different instruments. In Appendix B1.2 I derive regression weights for both instruments. Generally, because the past-cases-only instrument naturally relies more heavily on early cases, weighting schemes can differ if an examiner's leniency changes over time.

4 Data

4.1 Swedish Enforcement Authority

I use data on all applications to debt restructuring between 2010 and 2014 that made it through the initial screening provided by the SEA. The data contain an identifier for both the applicant and the respective examiner, the dates of initial application and the decision, as well as information on whether the application was accepted or rejected. For accepted applicants, I observe the level of debt as well as the overall amount to be paid back according to the initial repayment plan. I also observe if an application for an adjustment of the repayment plan was made and if this application was accepted or rejected. However, I do not observe the resulting changes in the repayment plan. I restrict my analysis to applicants of working age between 18 and 65. I also drop applicants whose debt is written off. The average acceptance rate is 61 percent.

4.2 Labor market outcomes and demographics

My data on labor market outcomes and demographics come from three administrative sources. The longitudinal integrated database for health insurance and labor market studies (LISA) contains information on labor market income, years and field of education, gender, age, unemployment insurance, and other forms of income such as pension income for all adult residents in Sweden. I use data for the years 2005 to 2021. The Structure of Earnings Survey is a survey containing all public sector employees and a random sample of around half of the private sector working population in Sweden. It is conducted by the Swedish National Mediation Office annually every September. This survey contains information about the number of hours worked and hourly wage rates. Finally, I use the Swedish linked employer-employee data from 2005-2019 to get information on individuals employers. I adjust all monetary variables to 2019 SEK prices. In the following main estimations, I keep a balanced sample such that I observe every applicant in LISA in all remaining seven years after the debt relief program begins. I further infer purchases of cars by ownership changes in the Swedish Vehicle Register.

4.3 Instruments and estimation sample

Appendix Table A8 shows average characteristics in the year before application for all applicants and split by acceptance status. The average applicant is 44.2 years old, has less than 11 years of education and has an annual labor income of 122,900 SEK which places them in the 23rd percentile of their birth cohort income ranking. Accepted applicants are less likely to be employed than rejected applicants (0.53 vs. 0.69), earn less (106,000 SEK vs. 148,800 SEK), and are more likely to be female (0.54 vs 0.43). The average amount of debt applicable for discharge is 761,400 SEK for accepted applicants which is 7.2 times the average income of accepted applicants. On average, they repay 104,900 SEK or 14.6 percent of their debt according to their initial repayment plan. 39.7 percent of accepted applicants do not have to repay anything according to their initial repayment plan.[18]

I construct two leniency instruments for each examiner. One based on the leave-one-out mean of all of the examiner's other decisions and one constructed as the average of past cases only.[19] To reduce noise I drop all observations that are estimated using 10 or less cases. It is common in the literature to first residualize each decision, for example, with the leave-one-out average decision within a court, as in Dobbie and Song (2015). Note that such a procedure can create a bias similar to the bias due to many controls as described in section 3.2.2 if there is path dependency in decision-making. Since randomization in my setting occurs on the national level, I do not need to make any similar adjustments. Appendix Figure A11 shows the distributions of both instruments. The average leave-one-out leniency is 0.63 and the average past leniency is 0.55. My final sample contains 185 examiners deciding on 99 cases on average. The leave-

[18] See Appendix Figure A10 for the distribution of the repayment rate.

[19] Since cases can overlap, I define past cases as all cases that are older than a month. My results are robust to using thresholds that are more restrictive.

one-out instrument for the average (median) observation is estimated using 346 (318) other cases. The past-cases-only instrument for the average (median) observation is estimated using 163 (110) cases.

4.3.1 Instrument relevance and validity

The independence assumption requires the instrument to be uncorrelated with the unobservable case characteristics the examiner bases her decision on. Failures of this assumption can arise from a lack of random allocation of examiners to cases, but also if path-dependency is a feature in the underlying decision-making process. Figure 1 plots the coefficients of OLS regressions of the two instruments on standardized applicant characteristics in the year prior to application, which we would expect to be predictive of U_{jt}.[20] The coefficients are small and insignificant for the past-cases-only instrument, but are substantially larger in magnitude and jointly highly significant for the leave-one-out instrument. These differences suggest that path dependency may be a relevant issue in this context.

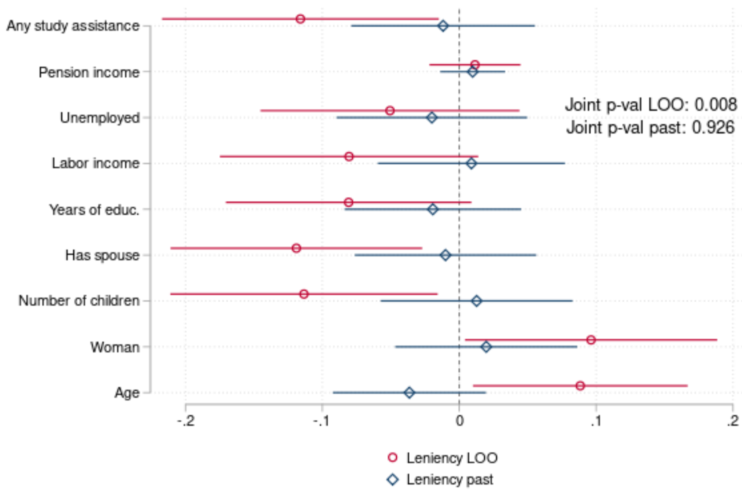


Figure 1: Balance of instruments

Note: This figures shows the coefficients of OLS regressions of the respective instrument Z_{jt}^{LOO} and Z_{jt}^{past} on standardized applicant characteristics defined in the year before application. The p-values refer to tests on joint significance from an OLS regression of the respective instruments on all characteristics. The balance variables are amount of study assistance, pension income, days unemployed, labor income, years of education, number of children below the age of 18 living in the same household, having a spouse (married or cohabiting), a female dummy, and age. Standard errors are clustered on the applicant level.

Table 1 presents the first stage regression of decisions on the two instruments. Both instruments are strongly correlated with the case decision, and the leave-one-out leniency IV appears to be the better predictor. The reason for this could be higher precision since leniency here is estimated from more cases than the past-cases-only instrument. An alternative or complementary explanation is that the leave-one-out instrument predicts more accurately by including

[20] I follow the recommendation in Chyn, Frandsen, and Leslie (2024) and cluster standard errors on applicant level only. A small share of around 6 percent of applicants applies more than once.

potentially endogenous information.

	(1) Accepted	(2) Accepted
Leniency LOO	0.804***	
	(0.0263)	
Leniency past		0.506***
		(0.0202)
Control for Year x Unit	Yes	Yes
F-Stat.	936	625
Observations	17032	17032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1: First stage

Note: This table shows coefficients and F-Statistics from regressions of the case decision on the instruments Z_{jt}^{LOO} and Z_{jt}^{past}. Standard errors are clustered on the applicant level.

The balance test in Figure 1 and the first stage estimation in Table 1 provide evidence for the independence assumption and the relevance assumption. The monotonicity assumption implies that an examiner who is strict with one applicant will also be stricter to all other applicants. I follow Bhuller, Dahl, Løken, and Mogstad (2020) and provide two tests for this assumption. First, if an examiner is strict towards one group, they should also be stricter towards other groups. I therefore split the sample by demographic characteristics and re-estimate the first stage within each of these sub-samples. If the monotonicity assumption holds, the leniency instrument should have the same sign in both sub-groups. Appendix Table A9 shows results for this test using the past-cases-only instrument. As required, all coefficients point in the same direction. Secondly, leniency estimates from one sub-sample should be positively correlated with the probability of acceptance in another sub-group. Appendix Table A10 tests this using the past-cases-only instrument. I split the sample by the same demographics as above, then I estimate leniency for each examiner in one of the splits, and regress the examiner's decision for all cases with applicants in the other sample on the leniency instrument obtained from the first sample. Again, all regressions show positive correlations.

5 Empirical results

5.1 The impact of debt restructuring on income and employment

Table 2 reports the main results on average income and average employment for the seven years following the start of the program. Columns 1 and 4 present the results for the respective OLS estimates. Accepted applicants earn on average 40,800 SEK less per year and are on average 15.4 percentage points less likely to be employed after the program started. Columns 3 and 6 show the 2SLS estimation results using the past-cases-only instrument. Both estimates drop considerably compared to the OLS estimation, suggesting negative selection into the program. Accepted applicants earn 14,400 SEK, or 9.1 percent, less per year than rejected applicants, but the estimate is insignificant. Employment is 9.2 percentage points (14 percent) lower as a direct effect of being accepted to the program.

	Income			Employed		
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	2SLS LOO	2SLS past	OLS	2SLS LOO	2SLS past
Accepted	-40766.8***	-25888.1*	-14389.9	-0.154***	-0.147***	-0.0917*
	(2408.4)	(11095.6)	(13310.0)	(0.00664)	(0.0313)	(0.0369)
Constant	175107.8***	165359.4***	158405.3***	0.702***	0.687***	0.653***
	(7581.7)	(6829.8)	(8172.0)	(0.0214)	(0.0192)	(0.0225)
Relative to control mean	-0.233	-0.157	-0.091	-0.220	-0.214	-0.140
Control for Year x Unit	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17032	17032	17032	17032	17032	17032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: OLS and 2SLS estimations

Note: This table shows coefficients for the OLS regressions and 2SLS regressions for each of the two instruments based on leave-one-out estimates of examiner leniency and past-cases-only. The outcomes are average annual labor income in SEK for 2019 prices and average of a dummy for having any employment in a given year for the seven years following the start of the debt relief program. I control for unit times year by using fixed effects in the OLS regression and by using a leave-examiner-out residualization in the 2SLS estimations. Standard errors are clustered on the applicant level.

5.1.1 Differences between leave-one-out and past-cases-only estimations

I report estimates for the leave-one-out instrument in columns 2 and 5. The estimate for income is -25,900 SEK and the estimate for employment is -14.7 percentage points, placing the two estimates between those of the OLS and the past-cases-only regressions. Two considerations are important when comparing the two instruments to each other. First, since both instruments are estimated from similar samples, they are highly correlated with each other. Using the standard errors of each regression to gauge the variation of their difference will, therefore, overestimate the respective standard error. Appendix Figure A12 plots bootstrap distributions of both estimators as well as their difference using a Bayesian or exchangeably weighted bootstrap (Rubin, 1989; Praestgaard and Wellner, 1993). The two panels show that both distributions of the differences $\hat{\beta}_{\text{past}} - \hat{\beta}_{\text{LOO}}$ have substantial mass above zero. The respective 95 percent confidence intervals are $[-4,719\ ;\ 29,099]$ for income and $[0.01\ ;\ 0.099]$ for employment. Secondly, both instruments and therefore their difference are related to the size of the OLS bias. In the structural model (3) this bias is scaled by the influence of the unobserved case characteristics U_{jt} on the outcome Y_{jt}, denoted by α. If α equals zero then both the OLS as well as the 2SLS regressions will in expectation yield the same estimate. To separate the question of how much the instruments differ from the question of how large the omitted variable bias in the OLS regression is, I scale the differences by the OLS bias. For this, I take the estimate of the past-cases-only IV as the ground truth and report the size of the leave-one-out IV bias compared to the OLS bias as

$(|\hat{\beta}_{\text{past}} - \hat{\beta}_{\text{LOO}}|)/(|\hat{\beta}_{\text{past}} - \hat{\beta}_{\text{OLS}}|)$. For income, I find a substantial difference of 43.6 percent compared to the OLS bias, for employment the relative difference is 88.8 percent.

As a sanity check that the timing of decisions drives the results, I construct an instrument using future cases only.[21] Appendix Table A12 shows the results for the OLS regression and the three 2SLS estimations.[22] Consistent with a model of path dependency in examiner decision-making, the bias of the future-cases instrument is substantially larger than the leave-one-out bias. For income, the future-cases estimation yields an estimate closer to the OLS estimate, for employment, it even exceeds the OLS estimate.[23]

Another potential reason why the two IV estimates might differ is differential weighting of underlying heterogeneous treatment effects. I estimate marginal treatment effects using a local IV approach (Heckman and Vytlacil, 2005; Brinch, Mogstad, and Wiswall, 2017) and report estimates of the local average treatment effect (LATE) and the average treatment effect (ATE) in Appendix Table A11. The ATE estimates are noisier since the underlying weights are not chosen based on efficiency considerations, but the difference between the leave-one-out and the past-cases-only point ATE estimates is even larger than for the LATE estimates. This suggests that the estimated bias of the leave-one-out estimation measured against the past-cases-only estimation is muted by the underlying weighting scheme.

In Appendix C1 I investigate the bias in leave-one-out examiner IV estimates in the setting of inventor mobility after patent approval similar to Melero, Palomeras, and Wehrheim (2020). I use the quasi-random allocation of patent applications to patent examiners within art unit at the US patent office to estimate whether granting a firm monopoly rights over the invention will change its incentives to retain inventors and therefore potentially hinder the diffusion of knowledge. I find a decrease of mobility of 3 percentage points using the leave-one-out instrument and a decrease of 6.6 percentage points using the past-cases-only instrument. This bias corresponds to 113 percent of the OLS bias, given an OLS estimate of -3.4.

5.1.2 Many controls bias

Next, I evaluate the bias caused by adding fixed effects. I randomly assign examiners to artificial fixed-effect groups. I then estimate both the 2SLS regression using the leave-one-out leniency IV and the past-cases-only leniency IV, including fixed effects, and benchmark these against the estimates of the past-cases-only IV using leave-examiner out residualization for each fixed effect group. I successively increase the number of fixed-effect groups from 2 to 92, so that each group contains around 2 examiners in the final estimation. I repeat this procedure 500 times for each fixed effect group size and report the absolute average difference between the

[21]Similar to the construction of the past-cases-only instrument, a future case is defined as one where the decision was made at least one month later.

[22]The estimates differ slightly to those reported in Table 2 since I further restrict the sample to observations for which the future cases instrument is estimated off more than 10 cases, similar to my treatment of the past cases instrument.

[23]Appendix Figure A13 provides further evidence for path-dependent decision making by showing that there is substantial autocorrelation in examiner decisions.

respective estimator and the residualized past-cases-only instrument. Panel (a) of Appendix Figure A14 plots the differences between the leave-one-out estimator with fixed effects and the leave-examiner-out residualized estimator using past-cases-only $|\hat{\beta}_{\text{LOO}} - \bar{\beta}_{\text{past}}|$ for the two main outcomes, income and employment. We can interpret this as the absolute bias of the leave-one-out estimation. The two biases from path dependency and many controls go into opposing directions.[24] Increasing the many control bias reduces the overall bias for both outcomes. When we compare the two estimators based on past-cases-only, including many fixed effects unambiguously increases the absolute bias. The corresponding results are plotted in panel (b) of Appendix Figure A14.[25]

5.1.3 Robustness

Appendix Tables A13 and A14 show results for different specifications of the past-cases-only estimation for income and employment, respectively. Column 1 shows the baseline specification with a balanced panel, controlling for unit-by-year, using only observations for which the instrument is constructed from at least 10 cases which have been decided 30 or more days before the previous case. In column 2, I allow for an unbalanced panel. Column 3 shows results without controlling for any covariates. Columns 4-6 restrict the sample to observations for which the instrument is constructed from at least 25, 75, and 125 cases, respectively. In column 7, I report estimates from instruments where past cases are defined as those which have been decided 60 or more days before. My main results are robust to all of these changes with point estimates for income being negative and insignificant and point estimates for employment being significant at around -10 percentage points or slightly larger. Appendix Tables A15 and A16 show the corresponding results using the leave-one-out instruments for income and employment, respectively. As in the baseline, these estimates are consistently larger in magnitude throughout all specifications. In my main specifications, I follow the literature[26] and do not adjust standard errors to account for the fact that the instrument itself is an estimation. In column 8, I instead use standard errors based on a Bayesian bootstrap that takes the construction of the instrument into account. Specifically, I estimate the instrument as a weighted average and then use the same weight for the respective observation in the 2SLS estimation.[27] The standard errors increase only slightly due to this adjustment.

[24]This can be seen from evaluating both biases individually as in (27) and (35). In Appendix B1.3 equation (36) is also explicitly derive the leave-one-out bias with many controls.

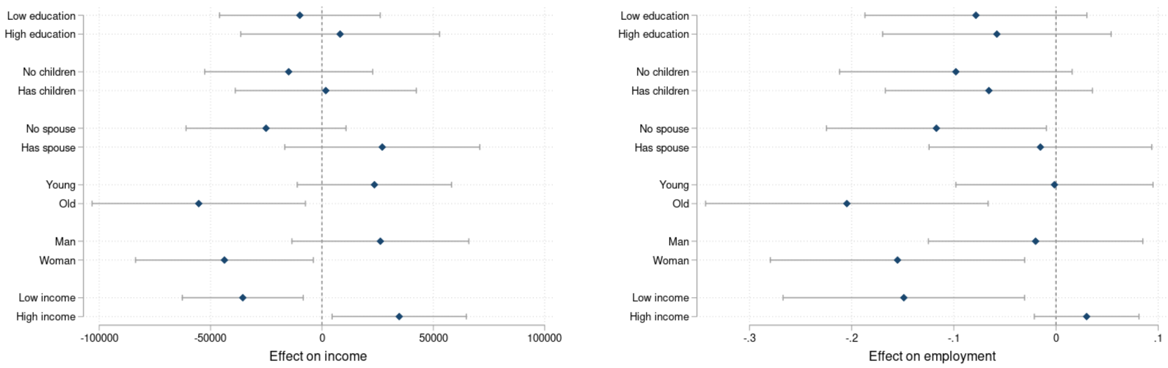
[25]Increasing the number of random fixed effect groups lower precision of the estimates. Appendix Figure A15 plots the first stage F-statistics for the three estimations over the number of fixed effect groups.

[26]See e.g. Autor and Houseman (2010), Maestas, Kathleen, and Strand (2013), Dahl et al. (2014), Dobbie and Song (2015), Melero et al. (2020), Bhuller et al. (2020), Gross and Baron (2022), Bruze et al. (2024), Collinson et al. (2024), Grindaker et al. (2024)

[27]I re-scale weights in the second step to account for changes in the sample due to balancing.

5.2 Heterogeneity and policy relevant treatment effects

The reported estimates on income and employment mask large heterogeneity. I split the sample by applicants' characteristics in the year prior to application to investigate these heterogeneities. Figure 2 reports the respective 2SLS coefficients using the past-cases-only instrument for income and employment. I split the sample by having above or below median years of education, age, and income[28] as well as having a child, having a cohabiting spouse, and gender. There are no large differences for applicants with low and high education as well as those who have or do not have children or a spouse. However, older applicants experience income losses of 55,300 SEK and a drop in employment of 20.5 percentage points while younger applicants' incomes increase by 23,600 SEK and their employment remains unchanged. Similarly, women do worse compared to men with treatment effects on income of -44,000 SEK and insignificant 26,240 SEK, respectively, and effects on employment of -15.5 percentage points for women and insignificant -0.2 percentage points for men. Differences in treatment effects are most pronounced for above-median and below-median income applicants. Relatively richer applicants' incomes increase by 34,700 SEK after being accepted to the program while they drop by 36,000 SEK for relatively poor applicants. Similarly, employment of above-median income participants increase by an insignificant 3 percentage points and decline by 14.9 percentage points for below median income participants.



(a) Income

(b) Employment

Figure 2: Heterogeneity by applicant characteristics

Note: These figures show the coefficients and 95 percent confidence intervals from estimating the 2SLS regression using the past-cases-only instrument. The sample is split by applicant characteristics measured in the year prior to application. The first split is by below and above median years of education, the second split is by having any children, the third split is by cohabiting with a spouse. The fourth split is by below and above median age, the fifth by gender, and the last split by having below and above income. The outcomes are average annual labor income in SEK for 2019 prices and the average of a dummy for having any employment in a given year for the seven years following the start of the debt restructuring program. Standard errors are clustered on the applicant level.

Another margin of heterogeneity can be explored by estimating marginal treatment effects (MTEs) for each level of examiner leniency. I follow Heckman and Vytlacil (2005) and estimate MTEs for each level of unobserved applicant resistance to treatment. This unobserved resistance

[28]The median applicant has 11 years of education, is 46 years old, and has an annual income of 31,233 SEK.

maps one-to-one into values of the residualized past-cases-only instrument.[29] Figure 3 plots the respective MTEs.[30] Treatment effects for the marginal applicant are negative for the most lenient examiners and grow with the strictness of the examiner. This is in line with the previous results split by income. Since applicants are evaluated by how likely it is that they will be able to repay their debt, applicants with lower income are more likely to be accepted to the debt restructuring program. The marginal applicant evaluated by a lenient examiner will on average have a lower income, than the marginal applicant evaluated by a strict examiner. The effect at the median leniency is zero.

The switch in sign of MTEs has implications for the overall effect of policy reforms that expand the debt restructuring program when abstracting from potential changes in the pool of applicants. We can view such an expansion as an increase in average examiner leniency. However, the overall effect will depend on where in the leniency distribution these changes occur. Examiners who are relatively lenient to begin with will accept further applicants who will as a result experience lower incomes. Examiners who are stricter will accept additional applicants whose incomes will increase as a result of undergoing debt restructuring.

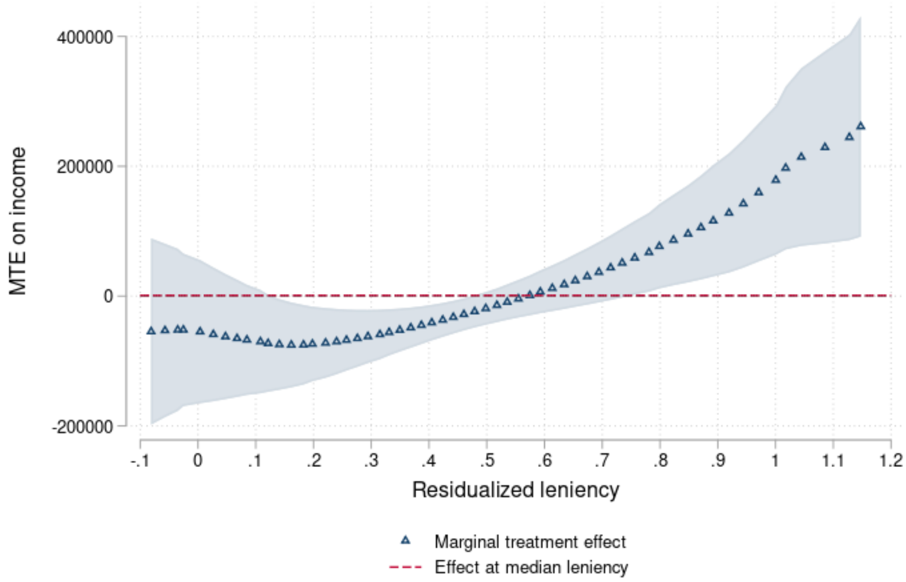


Figure 3: Marginal treatment effects by estimated leniency

Note: This figures plots the distribution of marginal treatment effects estimated on average income in the seven years after the program starts using a semi-parametric local IV estimation (Andresen, 2018) against estimated residualized leniency using past-cases-only. Leniency is residualized by year times office using a leave-examiner-out approach adding average leniency. The dotted red line marks the marginal treatment effect at the median leniency in the sample. 95 percent confidence intervals are computed by bootstrapping with 300 repetitions.

In Section 6 I will argue how the structure of the repayment plan can create different incentives based on initial employment, which can explain the heterogeneity in results (Appendix Fig-

[29] As in the previous estimations, I residualize leniency by office and year and then add the overall mean to the residual. leniency can therefore exceed 1 and be smaller than 0.

[30] In Appendix Figure A16, I report the corresponding estimates on employment.

ure A17 replicates the heterogeneity analysis of Figure 2 by further splitting the samples into above and blow-median income. This further split can explain a substantial part of the initial heterogeneity). In the following section I will therefore split the sample into tertiles by initial income.[31] Table 3 shows the respective 2SLS estimates for income and employment. Incomes of participants in the richest tertile increase by 44,235 SEK or 16.9 percent while incomes for participants in the middle and lowest tertiles drop by 48,474 SEK (35,9 percent) and 39,669 SEK (60 percent) relative to rejected applicants, respectively. Employment of high-income participants increases insignificantly by 3.9 percentage points (4.3 percent), whereas employment of middle and low-income participants decreases by 15.5 percentage points (23.3 percent) and 26.5 percentage points (63.2 percent), respectively.

	Income			Employed		
	(1)	(2)	(3)	(4)	(5)	(6)
	High income	Medium income	Low income	High income	Medium income	Low income
Accepted	44243.9*	-48473.9**	-39668.9*	0.0394	-0.153*	-0.265***
	(17417.6)	(18082.3)	(17788.8)	(0.0245)	(0.0655)	(0.0778)
Constant	262523.8***	135180.7***	66152.1***	0.915***	0.656***	0.420***
	(9947.3)	(11713.2)	(12687.9)	(0.0141)	(0.0424)	(0.0556)
Relative to control mean	0.169	-0.359	-0.600	0.043	-0.233	-0.632
Control for Year x Unit	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5402	5629	4716	5402	5629	4716

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: 2SLS estimations by income

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are average annual labor income in SEK for 2019 prices and the average of a dummy for having any employment in a given year for the seven years following the start of the debt restructuring program. I split the sample into tertiles using the average income in the two years prior to applications residualized by calendar year. I control for unit times year by using a leave-examiner-out residualization in the 2SLS estimation using past-cases-only. Standard errors are clustered on the applicant level.

5.3 Dynamic effects during and after the program

The outcomes of my main estimates refer to the seven years after the start of the debt restructuring program. In Tables 4 and 5 I show that the effects on income and employment persist even after the program has ended for both high and low-income applicants. High-income applicants experience an average increase of 36,730 SEK (14 percent) during their participation in the debt restructuring program and an increase of 58,580 SEK (22.3 percent) after the program ends relative to rejected applicants. The estimates on employment are 5.2 percentage points (5.7 percent) during the program and an insignificant 3 percentage points (3.3 percent) afterward. Incomes of low-income participants are 46,040 SEK (71.9 percent) lower than those of rejected applicants during the program and an insignificant 4,512 SEK (55.4 percent) lower after the program ends. The respective estimates for employment are -28.4 percentage points (68 percent) and -24.4 percentage points (56.6 percent).

[31] The average initial income in the lowest tertile is 37 SEK (only 1 percent of applicants have any income), in the second tertile is 22,700 SEK, and in the highest tertile is 284,200 SEK.

	High income		Low income	
	(1)	(2)	(3)	(4)
	During program	After program	During program	After program
Accepted	36725.2*	58579.9**	-46042.0**	-45119.3
	(16541.1)	(21814.1)	(15542.2)	(24653.5)
Constant	262836.6***	262563.2***	64079.5***	81396.8***
	(9453.8)	(12466.8)	(11121.0)	(17597.7)
Relative to control mean	0.140	0.223	-0.719	-0.554
Control for Year x Unit	Yes	Yes	Yes	Yes
Observations	5402	5402	4716	4716

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Income during and after the program

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are average annual income in SEK during the five years after the program began and for the two years after the program ended. I split the sample into tertiles using the average income in the two years prior to applications residualized by calendar year and report results for the highest and lowest group. I control for unit times year by using a leave-examiner-out residualization. Standard errors are clustered on the applicant level.

	High income		Low income	
	(1)	(2)	(3)	(4)
	During program	After program	During program	After program
Accepted	0.0523*	0.0299	-0.284***	-0.244**
	(0.0219)	(0.0364)	(0.0762)	(0.0944)
Constant	0.924***	0.896***	0.417***	0.431***
	(0.0127)	(0.0209)	(0.0544)	(0.0676)
Relative to control mean	0.057	0.033	-0.680	-0.566
Control for Year x Unit	Yes	Yes	Yes	Yes
Observations	5402	5402	4716	4716

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Employment during and after the program

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are average employment during the five years after the program began and for the two years after the program ended. I split the sample into tertiles using the average income in the two years prior to applications residualized by calendar year and report results for the highest and lowest group. I control for unit times year by using a leave-examiner-out residualization. Standard errors are clustered on the applicant level.

The precision of the IV estimations does not allow me to go further in time than seven years. In Figure 4, however, I provide graphical evidence indicating that divergences between accepted and rejected applicants persist up to nine years after the program started. I plot event study graphs comparing accepted applicants to a matched group of rejected applicants for applicants with initially high and initially low incomes.[32] The event studies suggest that income and employment of high-income participants increase directly after the start of the program and remain at an elevated level relative to the control group. For low-income participants employment drops with begin of the program while income steadily declines. After the program ends both employment and income increase slightly, but remain at substantially lower levels. In Appendix Figure A18, I plot for the raw means of income and employment for the full unmatched sample of accepted and rejected applicants.

To explore the debt restructuring program's effect on consumption dynamics I report estimates on the probability of buying a car during and after the program for high and low-income participants in Appendix Table A17. The estimates for low-income participants are insignificant both during and after the program. Similarly, high-income participants are neither more or less likely than rejected applicants to purchase a care while they are participating in the program. After the program ends, however, car purchases increase substantially by 21.2 percentage points or 165 percent. This can be indicative of generally higher consumption since debt repayments drop to zero after the program ends. It could also be a result of a relaxation of borrowing constraints. In both cases, these results suggest that high-income participants are able to improve their situation, while low-income participants are not.

6 Why does debt restructuring have adverse effects on some participants?

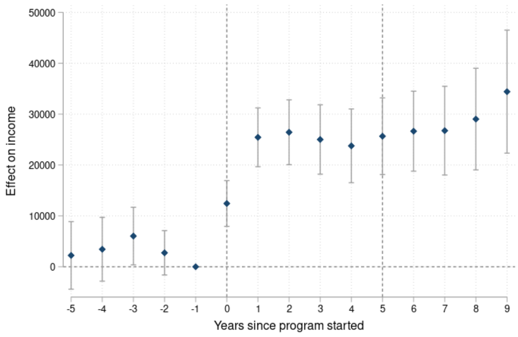
6.1 Wage garnishment and adjustable repayment plan

In the following section, I will discuss and provide evidence for mechanisms that explain the discrepancy in outcomes for high and low-income participants. By being accepted into the debt restructuring program, high-income applicants avoid wage garnishment.[33] Since the amount of wage garnishment is regularly adjusted based on changes in the debtor's income, it effectively functions like a tax on labor income and can therefore reduce incentives to work.[34] Further-

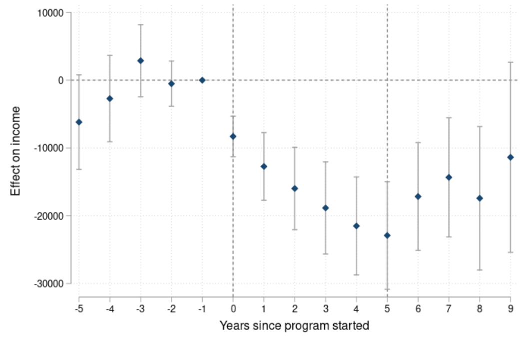
[32]I perform the matching by estimating propensity scores using a logit model and second-order polynomials of age and years of education in the year prior to application, a fully interacted set of dummies for gender, having a spouse, and having any children below the age of 18 in the same household, and income 3 years before application. I then predict propensity scores and generate 20 discrete bins. Within each bin and application year, I keep the same number of accepted and rejected applicants. The event-study is a regression of the outcome on event-time dummies interacted with a binary treatment indicator and event-time and applicant fixed effects. By matching on year of application and using event-time fixed effects, I circumvent issues due to treatment effect heterogeneity (Borusyak, Jaravel, and Spiess, 2024).

[33]See Appendix Figure A19 for the shares of accepted and rejected applicants subject to wage garnishment.

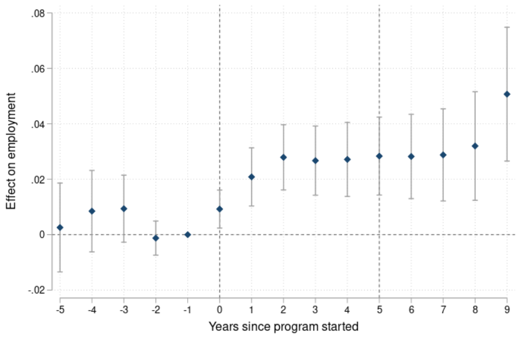
[34]This argument is also made by Dobbie and Song (2015) and Bruze et al. (2024). Dobbie and Song (2015) show that their results vary the strictness of state wage garnishment laws, but both papers do not have access



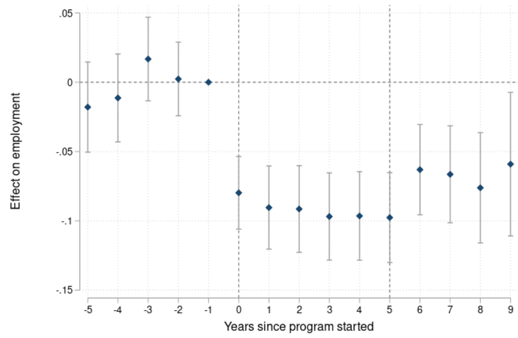
(a) Income - high income applicants



(b) Income - low income applicants



(c) Employment - high income applicants



(d) Employment - low income applicants

Figure 4: Event studies on income and employment

Note: These figures show the coefficients and 95 percent confidence intervals from estimating a dynamic difference-in-differences event study comparing accepted applicants to a matched sample of rejected applicants splitting the sample by initial income in the year before application. The matched group is chosen by estimating a logit propensity score model using second-order polynomials of age and years of education in the year prior to application, a fully interacted set of dummies for gender, having a spouse, and having any children below the age of 18 in the same household, and income 3 years prior to application. I predict propensity scores and generate 10 discrete bins. Then within each bin and application year, I keep the same number of accepted and rejected applicants. The outcome is annual income in SEK. I omit period $s = -1$ as the reference category. The first vertical dashed line indicates the start of the program and the second dashed line indicates the end after five years. Standard errors are clustered on the applicant level.

more, full-time employed participants are unlikely to face adjustments to their repayment plan since they are generally permitted to earn 4,000-5,000 SEK additionally per month and higher incomes resulting from normal wage increases are not considered as improvements in debtors' financial circumstances by the SEA. If the repayment due to the debt restructuring plan is considered as independent of labor income an additional income effect will further increase labor supply. Table 6 shows the effects on having wage garnishment and the average implied garnishment rate relative to income by initial income. While 27.6 percent of accepted high-income participants avoid wage garnishment, only 11.8 percent of low-income participants would have faced garnishment if they had been rejected. Additionally, the implied average garnishment rate for those who work is 6.9 percent for high-income participants and an insignificant 3.6 percent for low-income participants.

	Has garnishment			Garnishment rate		
	(1)	(2)	(3)	(4)	(5)	(6)
	High income	Medium income	Low income	High income	Medium income	Low income
Accepted	-0.276***	-0.140***	-0.118**	-0.0691***	-0.0398*	-0.0357
	(0.0364)	(0.0304)	(0.0395)	(0.0124)	(0.0176)	(0.0373)
Constant	0.311***	0.186***	0.164***	0.0762***	0.0554***	0.0578*
	(0.0212)	(0.0201)	(0.0284)	(0.00726)	(0.0108)	(0.0229)
Relative to control mean	-0.890	-0.753	-0.719	-0.906	-0.719	-0.617
Control for Year x Unit	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5402	5629	4716	5390	4209	1926

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Wage garnishment

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are the average probability of being subject to wage garnishment in the seven years after the program started and the average implied wage garnishment rate relative to income for those with positive income. I split the sample into tertiles using the average income in the two years prior to applications residualized by calendar year. I control for unit times year by using a leave-examiner-out residualization in the 2SLS estimation using past-cases-only. Standard errors are clustered on the applicant level.

To provide evidence that the effects on income increase with the likelihood of avoiding wage garnishment, I predict wage garnishment and estimate treatment effects separately for applicants with high and applicants with low predicted wage garnishment. I randomly split the sample of high-income applicants, using the first sub-sample to predict the probability of being subject to wage garnishment for the second sub-sample and vice versa. I then split the full sample again into a group with a below-median and an above-median predicted wage garnishment probability. Table 7 shows the 2SLS estimates on the probability of being subject to wage garnishment and income for the two sub-samples of high-income applicants. Columns 1 and 3 show that participants with higher predicted wage garnishment indeed experience a larger drop in their wage garnishment share compared to rejected applicants. Columns 2 and 3 show that the increase in income is substantially larger in the sample with high predicted garnishment, providing evidence that the counterfactual does shape the results for high-income applicants.[35]

to individual-level wage garnishment data.

[35] Due to the lower level of wage garnishment in the low-income group, the prediction exercise does not lead to different rates of actual wage garnishment. I report the corresponding results in Appendix Table A18.

	Low predicted garnishment		High predicted garnishment	
	(1)	(2)	(3)	(4)
	Has garnishment	Income	Has garnishment	Income
Accepted	-0.197***	26044.1	-0.314***	60656.5**
	(0.0580)	(28847.3)	(0.0477)	(21497.5)
Constant	0.256***	256038.5***	0.337***	271424.0***
	(0.0360)	(17626.9)	(0.0259)	(11446.9)
Relative to control mean	-0.767	0.102	-0.934	0.223
Control for Year x Unit	Yes	Yes	Yes	Yes
Observations	2696	2696	2695	2695

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: By predicted wage garnishment - high income

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are the average probability of being subject to wage garnishment in the seven years after the program started and average income measured in SEK in 2019 prices for high-income applicants. I split the sample and use the first sub-sample to predict the probability of being subject to wage garnishment out-of-sample in the second sub-sample and vice versa. To predict wage garnishment, I estimate an OLS regression of having wage garnishment on 5 bins of pre-application income and years of education, 10 bins of age in the year before application and dummies for gender, having a spouse, and having children as well as the year of application. Columns 1 and 3 show the actual share of individuals under wage garnishment for applicants with low and high predicted wage garnishment respectively, and columns 2 and 4 the effects on income measured in SEK in 2019 prices. Standard errors are clustered on the applicant level.

Unlike employed high-income participants, low-income participants can typically only increase their labor income discretely since most jobs require them to work a minimum amount of hours.[36] It is therefore highly likely that by finding a job, a participant will experience an increase in income above the adjustment threshold, which can result in an adjustment of the repayment plan. For incomes above the threshold, participants potentially face an effective marginal adjustment rate of 100 percent. Though creditors must actively apply for upward adjustments to the repayment plan, these rules are very salient to debtors, as they are prominently mentioned on the SEA's website and communicated by SEA staff. Figure 5 plots the probability that the repayment plan will be adjusted in the same or following year by changes in annual income compared to initial income for participants with no initial income. The average adjustment probability for participants whose income increase lies below the annual threshold of 48,000 SEK is around 1.7 percent. For participants whose income increases by more than 48,000 SEK the adjustment probability increases to 3.2 percent and further to 5.3 percent for those with large increases above 200,000 SEK. These results might indicate a very low risk of facing repayment plan adjustments. However, drawing conclusive evidence is challenging because we are conditioning on outcomes. For instance, the approximately 10 percent of participants who significantly increase their income may have agreements with their creditors preventing

[36] Appendix Figure A20 shows the distribution of weekly contracted hours from the Structure of Earnings Survey. Most employees work full-time at 40 hours per week. Only very few workers work less than 20 hours.

adjustments.[37]

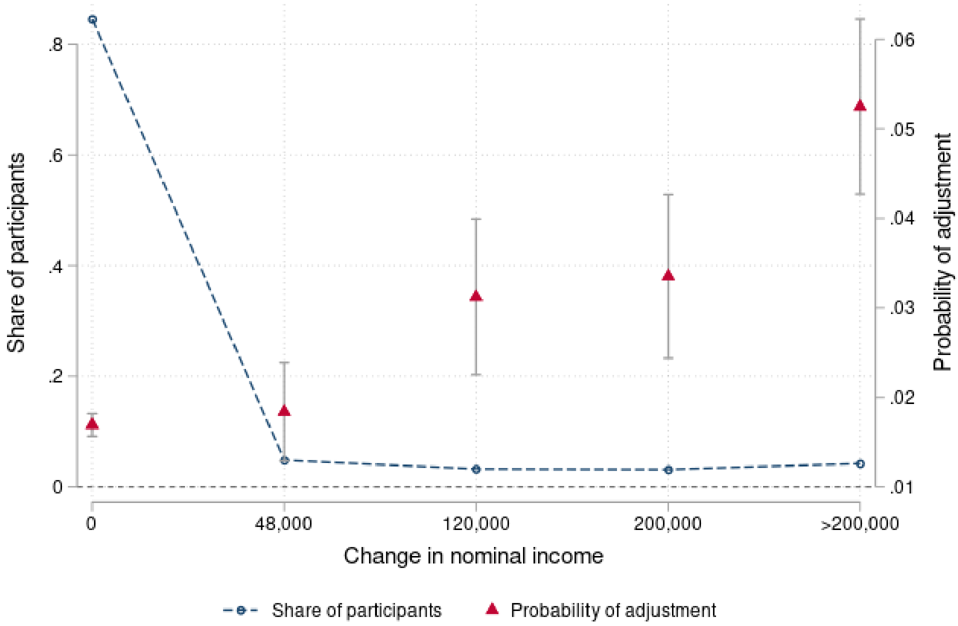


Figure 5: Adjustment by change in income for initially unemployed

Note: The red triangles show average rates of adjustments of the repayment plan in the same or following year for bins of changes in nominal income for participants without any initial labor income (right y-axis). The first bin includes everyone without any changes in income, the second bin includes participants whose income changes by less than 48,000 SEK, the third bin participants whose income increase between 48,000 and 120,000 SEK, the fourth bin participants whose income increase between 120,000 and 200,000 SEK, and the last bin all participants with income increases of more than 200,000 SEK. The blue dotted line shows the share of participants in the respective bin.

6.2 A model of labor supply under debt restructuring

In the following section, I develop and calibrate a model that highlights the mechanisms discussed above and allows me to evaluate counterfactual policies in which I vary the marginal adjustment rate that debtors with large increases in income face. I will abstract from applications of creditors for repayment plan adjustments and will assume that adjustments follow income changes deterministically given the salience of these rules for applicants to the debt restructuring program. The model generates persistence of the debt restructuring program's effects on participants by incorporating both wage growth through human capital accumulation by working as well as wage decreases caused by human capital depreciation as in Dinerstein et al. (2022). After five years without work, potential wages of participants will have declined

[37]Appendix Figure A21 plots a similar graph for participants who initially earn more than 160,000 SEK and are therefore likely to be full-time employed. Unlike in Figure 5 the first bin includes participants who experience a decrease in income and the second bin subsumes both small increases and participants without any change in incomes. The base rate of adjustments is higher at around 4.8 percent. This is likely driven by the fact that most initially unemployed participants are not required to repay anything which rules out downward adjustments. Compared to initially employed participants with low income changes the probability only increases significantly with very large income increases of more than 200,000 SEK. Similarly adjustments are more likely in cases where applicants experience a drop in incomes.

enough to prevent re-entry into the labor force. Another mechanism, which I abstract from, that could explain why participants remain unemployed after the program ends, could be signalling and discrimination by potential employers (Kroft et al., 2013; Eriksson and Rooth, 2014).

6.2.1 Model set-up

Individuals choose consumption c_t and weekly working hours h_t following MaCurdy (1981) preferences

$$u(c_t, h_t) = \sum_{t=0}^{T} \delta^t \left[\frac{c_t^{1-\gamma}}{1-\gamma} - \psi \frac{h_t^{1+\chi}}{1+\chi} \right] \tag{14}$$

where δ is a discount factor. I assume throughout that individuals cannot save or borrow.[38] The budget constraint is given by

$$c_t = w_t h_t - R_D^t(w_t h_t, y_r, r, h_r) \tag{15}$$

$$c_t \geq c_L. \tag{16}$$

w_t is the hourly net wage rate, R_D^t is the debt repayment schedule which I describe in detail below, and c_L is a subsistence minimum of consumption that individuals are guaranteed by the government in every period. Hours can be chosen from either non-employment or a discrete set of potential contracts which require at least 20 hours of work such that

$$h_t \in \{0, 20, 25, 30, 35, 40\}. \tag{17}$$

A minimum amount of hours can arise from either fixed costs of employment for potential workers, arising for example through commuting costs, or because of fixed costs for employers due to hiring or training new employees. I choose 20 hours as the minimum number of hours based on the distribution of contracted hours reported in the Structure of Earnings Survey (see Appendix Figure A20).

To explain the persistence of effects and to capture dynamic returns to working, I let wages follow the law of motion in Dinerstein et al. (2022)

$$w_{t+1} = (1 + \kappa_1 \mathbf{1}\{h_t > 0\} - \kappa_2) w_t. \tag{18}$$

κ_1 captures increases in human capital accumulated through work experience in the previous year and κ_2 captures human capital depreciation which arises in every year independently of an individual's work status.

[38] Borrowing constraints are a reasonable assumption for over-indebted individuals. Not allowing for savings and borrowing for individuals after they finish the debt restructuring program is inconsequential in this set-up, because consumption smoothing would require transferring income to earlier periods rather than saving during the repayment.

6.2.2 Repayment plan

I assume that rejected applicants to the debt restructuring program repay their debts through wage garnishment which I model as a tax on labor income. This tax is linear at rate τ if the remaining income is above the subsistence minimum c_L. Otherwise only the income exceeding c_L is taxed at τ.

$$R^t_{D=0} = \begin{cases} \tau w_t h_t & \text{if } (1-\tau) w_t h_t \geq c_L \\ \tau(w_t h_t - c_L) & \text{if } w_t h_t \geq c_L \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

For accepted applicants, repayment during the program depends on reference income y_r, reference hours h_r, and the initially set monthly repayment r. If monthly labor income does not exceed $y_0 + 4,000$ SEK or the additional income results from normal wage increases and not because the participant works more hours, the repayment will remain fixed at r. If the applicant increases their income by more than 4,000 SEK by working more, the additional income exceeding 4,000 SEK, will be taxed at the marginal adjustment rate $\rho = 1$.

$$R^{t<5}_{D=1}(w_t h_t, y_r, r, h) \begin{cases} r & \text{if } w_t h_t \leq y_r + 4000 \text{ or } h_t = h_r \\ \rho(w_t h_t - y_r - 4000) + r & \text{otherwise.} \end{cases} \tag{20}$$

Since income is purely a function of deliberate choices and there are no unforeseeable shocks to participants' financial capacities, the repayment plan cannot be adjusted downward. If a participant defaults on her repayments she exits the debt restructuring program and is subject to wage garnishment again. If she does not default, then the program ends after five years and all remaining debts are discharged

$$R^{t\geq 5}_{D=1} = 0. \tag{21}$$

6.2.3 Calibration and choice of parameters

I follow Keane and Wasi (2016) and set $\gamma = 0.727$. Kleven and Schultz (2014) estimate that the uncompensated elasticity of earned income with respect to a large tax change in Denmark is 0.26. I set $\chi = 0.325$ so that the uncompensated labor supply elasticity of the static model given γ matches the estimate in Kleven and Schultz (2014).[39] I set the discount factor $\delta = 0.94$. The experience effect on human capital $\kappa_1 = 0.068$ and human capital depreciation $\kappa_2 = 0.043$ following the estimates in Dinerstein et al. (2022). The average employed participant earns a gross hourly wage of 148 SEK. I abstract from non-linearities of the Swedish tax schedule and deduct a linear income tax of 14 percent.[40] I set $\tau = 0.12$ to the median wage garnishment

[39] The uncompensated elasticity of the static problem is given by $\epsilon^L_U = \frac{1-\gamma}{\chi+\gamma}$, see Keane (2011).

[40] In reality, the Swedish income tax schedule is non-linear due to an earned income tax credit and deduction from taxable income and can vary depending on the municipality of residence. 14 percent is a reasonable approximation of the average tax payed by either part-time or full-time workers with an hourly gross wage of 148 SEK.

rate prior to the program. I then calibrate the disutility of labor $\psi = 0.02$ to match net wages and average hours of 35 hours per week of participants in the year before application.[41] I set monthly subsistence consumption to $c_L = 7,300$ SEK. Finally, I match the median of the empirical repayment rates and set initial repayment to 12 percent of initial income for those employed and to 0 for those without initial employment.

6.2.4 Changing the adjustment rate ρ

Figure 6 plots weekly hours for the year before application which I exogenously set weekly hours worked to 35 and the seven following years (see Appendix Figure A22 for monthly income). I consider two cases to evaluate whether the model can replicate my empirical findings. The orange dotted line plots hours for a rejected applicant who faces wage garnishment. Since the economic incentives for this individual do not change, labor supply remains at 35 hours throughout the entire period. The dashed blue line plots hours for a successful applicant who starts the debt restructuring program in year 0. This successful applicant increases labor supply by 5 hours per week throughout the five years of debt restructuring and the two subsequent years. Increasing labor supply by 5 hours per week is insufficient to trigger an adjustment of the repayment plan which is therefore fixed from the perspective of the participant. Therefore, the effective wage increases for participants and the income effect induced by the fixed repayment both increase incentives to work. However, the incentives induced by the higher effective wage are sufficient to maintain employment at 40 hours even after the program has ended. In column 1 of Appendix Table A19, I report a 2SLS estimate on intensive margin hours of 3.3 hours per week for high-income applicants.

The third, solid blue line in Figure 6 plots hours for an accepted applicant with the same features as the previous accepted applicant except that I exogenously set initial hours to zero. The resulting labor supply differs substantially, as the low-income participant remains without work throughout the entire duration of the program and the two years thereafter. During the program, potential adjustments of the repayment plan prevent this participant from seeking employment. At the same time, her productivity falls by 4.3 percent every year due to human capital depreciation. Unlike the employed participant, the unemployed participant does not experience any increases in human capital through work experience. The gross wage drops from 148 SEK to 113.7 SEK after six years out of the labor force. This is enough to prevent her from seeking re-employment after the program ends.

Next, I simulate policy reforms which reduce the marginal adjustment rate ρ and evaluate the resulting changes in labor supply of the initially unemployed accepted applicant. Figure A23 plots weekly hours worked for different values of ρ. If ρ decreases by 10 percentage points to 0.9, labor supply remains unaffected. However, already a decrease to 0.8 is sufficient to trigger a return to the labor market, with weekly hours increasing to 25 hours per week during the

[41] Specifically, I solve for optimal hours in the static setting and set $\psi = ((1-\tau)\bar{w}_N)^{1-\gamma}\bar{h}^{-\gamma-\chi}$ where $\bar{w}_N$ is the average net wage and $\bar{h}$ are the average monthly hours.

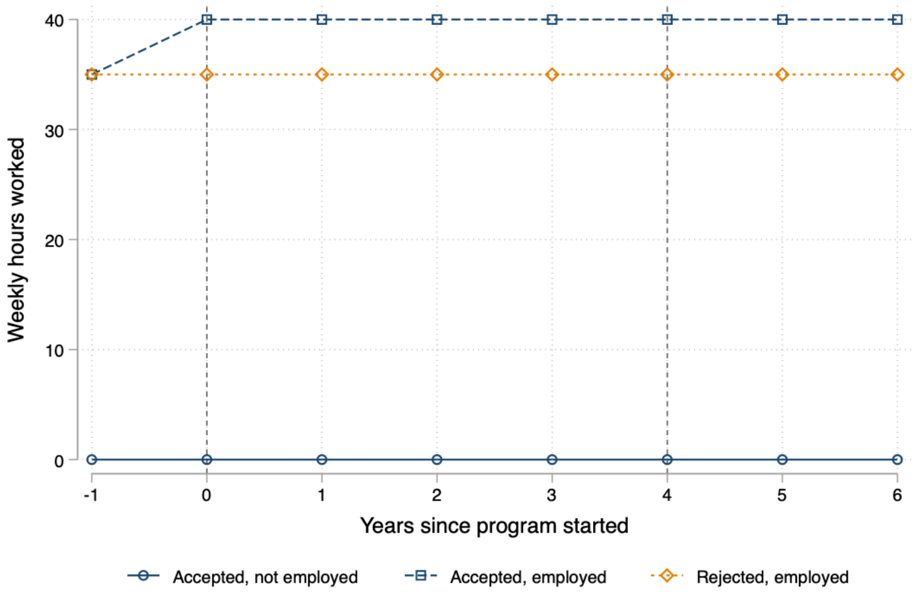


Figure 6: Weekly hours - model

Note: The graphs show weekly hours predicted by the model. Initial hours in period -1 are set endogenously. The dotted orange line plots outcomes for a rejected applicant who initially works 35 hours per week and who faces wage garnishment. The dashed blue line plots outcomes for an accepted applicant who initially works 35 hours per week. And the solid blue line plots outcomes for an accepted applicant who does not work in the year prior to application. The first vertical dashed line indicated the begin of the debt restructuring program and the second dashed line its end.

program and to 35 hours per week after the program ends and any disincentives induced by the adjustable repayment plan disappear.[42] For a rate $\rho = 0.2$ labor supply during the program increases to 30 hours per week. For adjustment rates less than or equal to 0.1 labor supply increases to 35 hours per week.

The additional value generated by the debtor's re-entry into the labor market is split among the debtor herself, her creditors as well as the government who pays out less in social benefits and collects more in income taxes. Figure 8 shows how the additional value over seven years is distributed for different values of the adjustment rate ρ. For a 20 percentage points lower adjustment rate, debtors enjoy additional overall consumption of almost 600,000 SEK. For a fixed repayment plan ($\rho = 0$), additional consumption increases to over 931,000 SEK compared to the case of $\rho = 1$. The additional debt repayment to creditors follows a Laffer-curve with a modest maximum value of 42,000 SEK at $\rho = 0.8$, but is positive for all ρ between 0.8 and 0.1. Government revenue increases by between 815,000 SEK and 865,000 SEK for adjustment rates smaller or equal to 0.8. The overall total value over seven years of reducing the adjustment rate by only 20 percentage points amounts to 1,439,000 SEK per initially unemployed applicant. This corresponds to 5.1 times the annual full-time gross salary at an hourly wage rate of 148.

Changes in the adjustment rate may not only result in different incentives for the initially

[42]Labor supply does not increase to 40 hours as for initially employed participants, since the initial year of unemployment results in a 6.8 percent lower wage due to forgone work experience.

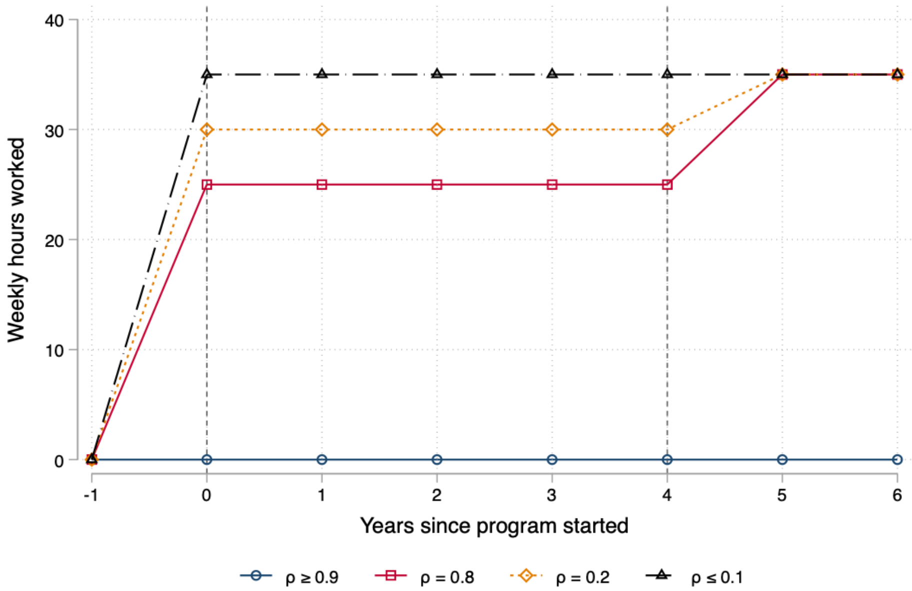


Figure 7: Weekly hours by marginal adjustment rate

Note: The graphs show weekly hours predicted by the model for different marginal adjustment rates ρ as in equation (20). All outcomes are for accepted applicants who are initially unemployed. The first vertical dashed line indicates the begin of the debt restructuring program and the second dashed line its end.

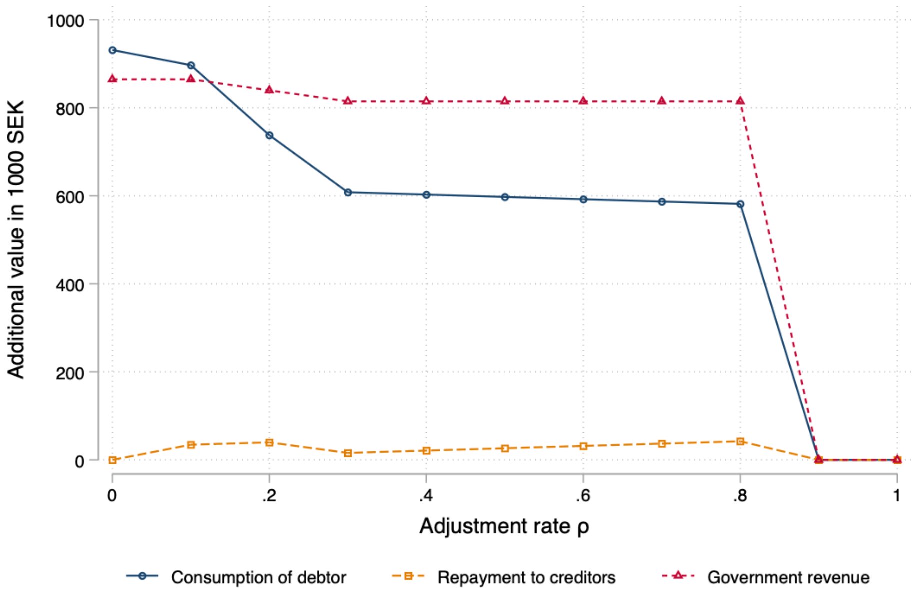


Figure 8: Additional monetary value by marginal adjustment rate

Note: The graphs additional value in 1000 SEK over seven years for different marginal adjustment rates ρ as in equation (20). The blue solid line plots additional consumption for debtors, the dashed orange line plots additional repayment to creditors, and the dashed red line additional government revenue stemming from lower social benefits payed out and higher labor taxes. All outcomes are for accepted applicants who are initially unemployed. The vertical dashed line indicates that all values to the left might not be incentive-compatible since applicants could prefer not to work when they apply for debt restructuring compared to working 35 hours.

unemployed, but can also lead to moral hazard for employed applicants. If the repayment plan is based on initial income and is either not adjusted or only minimally adjusted in response to increases in income, then this can create incentives for applicants to lower their debt repayment by working less or not at all before applying. I benchmark the expected utility of working fewer than 35 hours per week against the expected utility of working 35 hours and define expected utility as

$$E[u\,(h_{-1}; D, \rho)] = p \times u\,(h_{-1}; D = 1, \rho) + (1 - p) \times u\,(h_{-1}; D = 0, \rho) \tag{22}$$

where $p = \Pr(D = 1)$ is the probability of being accepted to the debt restructuring program. The relationship between initial income and probability of acceptance is not straightforward. On the one side, applicants are more likely to be accepted if the ability to repay their debt is smaller and hence lower initial incomes can increase p. On the other side, applicants are expected to have made their best effort to repay. Voluntarily decreasing labor supply might harm an applicant's chances of being accepted. Figure 9 therefore plots expected utility for different values of initial hours and adjustment rates ρ for four values of acceptance probability ranging from 0.2 to 0.8.[43] For almost all scenarios, applicants are better off working at least 30 hours per week. Only relatively high acceptance probabilities combined with very low adjustment rates could induce applicants to strategically exit the work force to lower repayment obligations. Depending on the probability of acceptance, applicants prefer to either work 30 or 35 hours per week. However, these differences are small and almost invariant to adjustment rates ρ. These results suggest that lowering adjustment rates is unlikely to result in additional labor supply moral hazard among applicants. A distinct issue that may be of concern is debtor moral hazard, or the amount of additional debt that households accumulate given the option of discharge through the debt restructuring program. While a thorough evaluation is beyond the scope of this paper, this second type of moral hazard is likely to be substantially more sensitive to changes in the acceptance probability faced by debtors rather than to changes in adjustment rates.

6.3 Comparison to debt relief programs in US and Denmark

To understand the disparities in outcomes between the Swedish debt restructuring program and those of the US and Denmark we have to consider both differences in the respective populations underlying the reported estimates as well as institutional details. Dobbie and Song (2015) find that in the US, Chapter 13 bankruptcy increases annual earnings by 25.1 percent and employment by 6.8 percentage points. These effects are driven by a drop in earnings and employment of rejected applicants who potentially face wage garnishment and home foreclosure. Chapter 13 bankruptcy allows the discharge of debt following a three to five year partial repayment period. As in the Swedish setting, the repayment plan can be adjusted if debtors see a significant increase in their income. Why do outcomes between Swedish and US debtors differ

[43]Note that while the scale of expected utility is irrelevant, the relative differences between different initial hours indicate how easily an applicant is shifted between states.

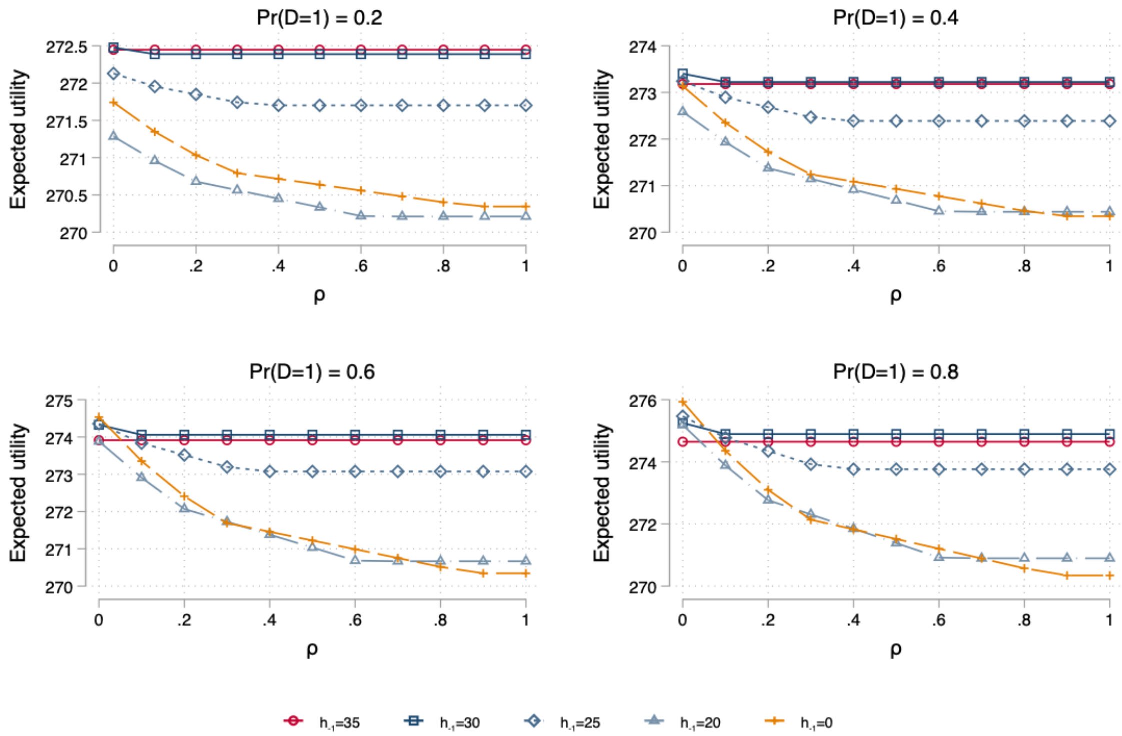


Figure 9: Expected utility for different adjustment rates and different probabilities of acceptance

Note: Each of the four panels plot expected utility of different values of initial weekly hours and different adjustment rates ρ as in equation (20). Expected utility is defined over the two states of being accepted too the debt restructuring program and being rejected. The four panels differ in the assumed probability of acceptances which is given by 20 percent in the upper left panel, 40 percent in the upper right panel, 60 percent in the lower left panel, and 80 percent in the lower right panel.

so much? A plausible explanation is the substantial difference in the underlying population of applicants. In the US, the average applicant has a 81.3 percent probability of being employed compared to 58 percent in Sweden. In a heterogeneity analysis, Dobbie and Song (2015) split the sample by initial median income.[44] Comparing high-income and low-income applicants, the effect in earnings decreases from 8,650 US dollars to 1,691 dollars and the effect on employment falls from 9.9 percentage points to an insignificant 3.9 percentage points. However, even low-income applicants in this setting have a 6.8 percentage point higher probability of being initially employed compared to applicants in the the overall Swedish sample. Investigating how these results change when considering only applicants without employment, who can increase their income only by large discrete steps and risk adjustments in their repayment plan, could shed light on the potential adverse effects of Chapter 13 bankruptcy on this vulnerable subgroup.

In Denmark, the debt relief program also allows for the discharge of unsecured debt after a partial repayment period of typically five years. Importantly, however, the repayment plan is fixed at the initial decision and does not vary with debtors' incomes. Bruze et al. (2024) report that participants in the Danish debt relief program have 26 percent higher earned income and are 11.7 percentage points more likely to be employed in the 16 years after the program starts. Danish applicants have a 64 percent probability of being initially employed, making them worse off than their US counterparts but better off than Swedish applicants. Notably, the difference in employment between accepted and rejected applicants in Denmark is -2.4 percentage points compared to -17 percentage points in Sweden. Bruze et al. (2024) also split the sample by initial median income.[45] Unlike in the US and Sweden, low-income applicants experience larger increases in labor earnings and employment hinting that there are no inherent reasons for why low-income individuals would react worse to debt relief.

[44] See Table 5 in Dobbie and Song (2015)

[45] See Appendix Table A.24 in Bruze et al. (2024). The authors do not provide summary statistics on employment or income of these two groups.

7 Conclusion

This paper analyzes whether the Swedish debt restructuring program is able to improve the economic conditions of its beneficiaries. While previous work in the US and Denmark finds overwhelmingly positive effects on later labor market and financial outcomes, I report negative average effects on labor market income and employment. Applicants with relatively high initial income are able to improve their economic conditions, whereas participants in the lower two-thirds of the initial income distribution experience large declines in both income and employment compared to rejected applicants. The changes induced by the program persist even after it has ended and can be explained by the interaction of wage garnishment as a counterfactual to debt restructuring and high potential adjustment rates in the repayment plan schedule for initially unemployed participants. I calibrate a model of dynamic labor supply and show that already modest changes to how repayment plans are adjusted to higher incomes can substantially improve debtor welfare and slightly increase repayments to creditors. Overall, my findings highlight that seemingly small differences in the institutional details of debt restructuring programs can result in vastly different outcomes for the most economically vulnerable groups of debtors. Given the rise in popularity of partial repayment debt relief programs worldwide, policy makers should pay attention to these details when designing policies.

A second contribution of this paper is to show that classical examiner or judge instrumental variable designs based on jackknife or leave-one-out estimates of examiner leniency will be biased under path-dependent decision-making when examiners current decisions are influenced by past case characteristics. This issue can be addressed by constructing leniency instruments based solely on past cases. I show empirical evidence for the presence of such past dependencies in decision-making in the context of debt restructuring in Sweden. I report that instruments based on either past-cases-only and future-cases-only yield different results in 2SLS regression from both each other and the standard leave-one-out approach.

Bibliography

Andresen, M. (2018). Exploring marginal tretament effects: Flexible estimation using stata. *The Stata Journal 18*(1), 118–158.

Angrist, J. D., G. W. Imbens, and A. B. Kruger (1999). Jackknife instrumental variables estimation. *Journal of Applied Econometrics 14*(1), 57–67.

Autor, D. H. and S. N. Houseman (2010). Do temporary-help jobs improve labor market outcomes for low-skilled workers? evidence from "work first". *American Economic Journal: Applied Economics 2*(3), 96–128.

Bekker, P. (1994). Alternative approximations of the distributions of instrumental variable estimators. *Econometrica 62*(3), 657–681.

Bergthaler, W., J. Garrido, and A. Rosha (2023). The right tool for the job? mortgage distress and personal insolvency during the european debt crisis. *IMF Working Paper No.2023/092*.

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020). Incarceration, recidivism, and emplyoment. *The Journal of Political Economy 128*(4), 1269–1324.

Bhuller, M. and S. Henrik (2024). Feedback and learning: The causal effects of reversals on judical decision-making. *Working Paper*.

Bindler, A. and R. Hjalmarsson (2019). Path dependency in jury decision making. *Journal of the European Economic Association 17*(6), 1971–2017.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting event study designs: Robust and efficient estimation. *Working Paper*.

Breitung, J., S. Kripfganz, and K. Hayakawa (2022). Bias-corrected method of moments estimators for dynamic panel data models. *Econometrics and Statistics 24*, 116–132.

Brinch, C., M. Mogstad, and M. Wiswall (2017). Beyond late with a discrete instrument. *Journal of Political Economy 125*(4), 985–1039.

Bruze, G., A. K. Hilsløv, and J. Maibom (2024). The long-run effect of individual debt relief. *Working Paper*.

Chao, J. C., N. R. Swanson, and T. Woutersen (2023). Jackknife estimation if a cluster-sample iv regression model with many weak instruments. *Journal of Econometrics 235*(18), 1747–1769.

Chen, D. L., T. J. Moskowitz, and K. Shue (2016). Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics 131*(3), 1181–1242.

Chyn, E., B. Frandsen, and E. C. Leslie (2024). Examiner and judge designs in economics: A practioner's guide. *Working Paper*.

Collinson, R., J. E. Humphries, N. Mader, D. Reed, D. Tannenbaum, and W. van Dijk (2024). Evicition and poverty in american cities. *The Quarterley Journal of Economics 139*(1), 57–120.

Dahl, G., A. Kostøl, and M. Mogstad (2014). Family welfare cultures. *The Quarterly Journal of Economics 129*(4), 1711–1752.

de Chaisemartain, C. and D'Haultfœuille (2020). Two-way fixed effects estimators with heterogenous treatment effects. *American Economic Review 110*(9), 2964–2996.

Dinerstein, M., M. Rigissa, and C. Yannelis (2022). Human capital depreciation and returns to experience. *American Economic Review 112*(11), 3725–3762.

Dobbie, W., J. Goldin, and C. Yang (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review 108*(2), 201–240.

Dobbie, W. and J. Song (2015). Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection. *American Economic Review 99*(5), 1272–1311.

Dohmen, T. and J. Sauermann (2016). Referee bias. *Journal of Economic Surveys 30*(4), 679–695.

Englich, B., T. Mussweiler, and F. Strack (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin 32*(3), 188–200.

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

Frandesen, B., E. Leslie, and S. Mcintryie (2023). Cluster jackknife instrumental variables estimation. *Working Paper*.

Galasso, A. and M. Schankerman (2015). Patents and cumulative innovation: Causal evidence from the courts. *The Quarterly Journal of Economics 130*(1), 317–370.

Gaulé, P. (2018). Patents and the success of venture-capital backed startups: Using examiner assignment to estimate causal effects. *The Journal of Industrial Economics 66*(2), 350–376.

Grindaker, M., A. Kostøl, and M. Merkle (2024). Layoff costs and the value of an employer to employees. *Working Paper*.

Gross, M. and E. J. Baron (2022). Temporary stays and persistent gains: The causal effects of foster care. *American Economic Journal: Applied Economics 14*(2), 170–99.

Heckman, J. and E. Vytlacil (2005). Structural equations, treatment effects, and economic policy evaluation. *Econometrica 73*(3), 669–738.

Imbens, G. W. and J. D. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Indarte, S. (2023). Moral hazard versus liquidity in household bankruptcy. *The Journal of Finance 78*(5), 2421–2464.

Jin, L., R. Tang, H. Ye, J. Yi, and S. Zhong (2023). Path dependency in physician decision making. *The Review of Economic Studies*.

Keane, M. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

Keane, M. and N. Wasi (2016). Labour supply: The roles of human capoital and the extensive margin. *The Economic Journal 126*(592), 578–617.

Kleven, H. and E. Schultz (2014). Estimating taxable income responses using danish tax reforms. *American Economic Journal: Economic Policy 6*(4), 271–301.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review 96*(3), 863–876.

Kolesár, M. (2013). Estimation in an instrumental variables model with treatment effect heterogeneity. *Working Paper*.

Kramer, R. S. S., A. L. Jones, and D. Sharma (2013). Sequential effects in judgements of attractiveness: The influences of face race and sex. *PLoS One 8*.

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

Larsson, B. and B. Jacobsson (2013). Discretion in the "backyard of law": Case handling of debt relief in sweden. *Professions & Professionalism 4*(1), 438–454.

Lennander, G. (1991). Debt adjustment for private individuals. *5 Stockholm Institute for Scandinavian Law 1129*.

Liu, Y. and C. B. Rosenberg (2013). Dealing with private debt distress in the wake of the european financial crisis a review of the economics and legal toolbox. *IMF Working Paper No.2013/044*.

MaCurdy, T. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy 89*(6), 1059–1085.

Maestas, N., M. Kathleen, and A. Strand (2013). Does disability insurance receipt discourage work? using examiner assignment to estimate causal effects of ssdi receipt. *American Economic Review 103*(5), 1797–1828.

Melero, E., N. Palomeras, and D. Wehrheim (2020). The effect of patent protection on inventor mobility. *Management Science 66*(12), 5485–5504.

Page, L. and K. Page (2010). Last shall be first: A field study of biases in sequential performance evaluation on the idol series. *Journal of Economic Behavior & Organization 73*(2), 186–198.

Praestgaard, J. and J. A. Wellner (1993). Exchangeably weighted bootstraps of the general empirical process. *The Annals of Probability 21*(4), 2053–2086.

Radbruch, J. and A. Schiprowski (2023). Interview sequences and the formation if subjective assessments. *Working Paper*.

Ramsay, I. (2017a). Personal insolvency in the 21st century. a comparative analysis of the us and europe. *Bloombury Publishing*.

Ramsay, I. (2017b). Towards an international paradigm of personal insolvency law? a critical view. *QUT Law Review 17*(1), 15–39.

Rubin, D. (1989). The bayesian bootstrap. *The Annals of Statistics 9*(1), 130–134.

Sampat, B. and H. L. Williams (2019). How do patents affect follow-on innovation? evidence from the human genome. *American Economic Review 109*(1), 203–236.

Srinivasan, K. (2023). Judical scarring. *Working Paper*.

Stavins, J. (2000). Credit card borrowing, delinquency, and personal bankruptcy. *Federal Reserve Bank of Boston New England Economic Review*, 15–30.

Appendix

A1 Tables

	All	Accepted	Rejected
Age	44.20	45.11	42.80
	(8.02)	(7.53)	(8.52)
Share women	0.50	0.54	0.43
	(0.50)	(0.50)	(0.50)
Years of education	10.84	10.73	11.00
	(1.90)	(1.85)	(1.96)
Number of children	0.70	0.67	0.74
	(1.10)	(1.08)	(1.12)
Labor income in 100 SEK	1229	1060	1488
	(1427)	(1356)	(1493)
Employed	0.59	0.53	0.69
	(0.49)	(0.50)	(0.46)
Within cohort income rank	0.23	0.20	0.28
	(0.24)	(0.22)	(0.25)
Debt in 100 SEK	-	7614	-
		(10183)	
Repayment in 100 SEK	-	1049	-
		(1417)	
Observations	17032	10301	6731

Table A8: Applicant characteristics

Note: This table shows means and standard deviations (in parentheses) for characteristics in the year before application for all applicants, those accepted, and those rejected. The characteristics are age, binary gender, number of children below the age of 18 living in the same household, labor income in 100 SEK measured in 2019 prices, a dummy for being employed, and the within birth cohort labor income rank. The total amount of debt in 100 SEK and the overall repayment according to the initial repayment plan in 100 SEK are only observed for accepted applicants.

	Split I		Split II		Split III		Split IV	
	(1) Women	(2) Men	(3) Young	(4) Old	(5) No College	(6) College	(7) Low Income	(8) High Income
Leniency past	0.518***	0.546***	0.685***	0.478***	0.528***	0.599***	0.460***	0.588***
	(0.0276)	(0.0279)	(0.0355)	(0.0233)	(0.0208)	(0.0626)	(0.0281)	(0.0271)
Observations	8503	8529	5122	11910	15323	1624	8195	8837

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A9: Test for monotonicity - sub-samples

Note: This figures shows the first stage coefficients and F-statistics for regressions of the examiner's decision on the past-cases-only instrument for different splits of the sample. The first two columns are split by gender. The third and fourth column are split into applicants of age 40 or younger and applicants above the age of 40. The fifth and sixth split is done by applicants with and without a college degree. The final split splits the sample by income into those with annual income smaller that 3,120 SEK and those above.

	Split I		Split II		Split III		Split IV	
	(1) Women	(2) Men	(3) Young	(4) Old	(5) No College	(6) College	(7) Low Income	(8) High Income
Leniency past	0.293***	0.426***	0.343***	0.287***	0.380***	0.251***	0.248***	0.414***
	(0.0223)	(0.0232)	(0.0255)	(0.0207)	(0.0212)	(0.0429)	(0.0219)	(0.0230)
Observations	9645	9648	5886	12433	14164	1866	9160	10165

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10: Test for monotonicity - out-of-sample

Note: This figures shows the first stage coefficients and F-statistics for regressions of the examiner's decision on the past-cases-only instrument. In each split I estimate leniency for one of the two sub-samples and then use this leniency on the other sub-sample. The first two columns are split by gender. The third and fourth column are split into applicants of age 40 or younger and applicants above the age of 40. The fifth and sixth split is done by applicants with and without a college degree. The final split splits the sample by income into those with annual income smaller that 3,120 SEK and those above.

	Income		Employed	
	(1) LOO	(2) Past	(3) LOO	(4) Past
effects				
ATE	-24535.0	13906.9	-0.104	-0.00260
	(25123.0)	(20704.2)	(0.0682)	(0.0512)
LATE	-21936.7*	-8241.4	-0.131***	-0.0728*
	(10716.9)	(13407.7)	(0.0319)	(0.0357)
Control for Year x Unit	Yes	Yes	Yes	Yes
Observations	17032	17032	17032	17032

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11: Average treatment effects

Note: This table shows estimates of the LATE and the ATE for the leave-one-out instrument and the instrument using past-cases-only using a semi-parametric local IV marginal treatment effect estimation (Andresen, 2018). The outcomes are annual labor income in SEK for 2019 prices and having any employment in a given year averaged over the seven years following the start of the debt restructuring program. I control for unit times year by using a leave-examiner-out residualization in the 2SLS estimation using past-cases-only and future cases. Standard errors are clustered on the applicant level.

	Income				Employed			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	2SLS LOO	2SLS future	2SLS past	OLS	2SLS LOO	2SLS future	2SLS past
Accepted	-41395.1***	-20444.0	-31369.7*	-8298.7	-0.156***	-0.138***	-0.186***	-0.0813*
	(2449.9)	(11376.4)	(14051.8)	(13705.1)	(0.00675)	(0.0322)	(0.0404)	(0.0379)
Constant	173614.2***	161787.3***	168409.5***	154425.9***	0.699***	0.681***	0.710***	0.646***
	(7945.5)	(7010.9)	(8610.3)	(8426.3)	(0.0224)	(0.0198)	(0.0247)	(0.0232)
Control for Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16459	16459	16459	16459	16459	16459	16459	16459

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12: Future cases IV

Note: This table shows coefficients for the OLS regressions and 2SLS regressions for the leave-one-out instrument, the instrument using past-cases-only, and an instrument using only future cases. I restrict the samples to observations for which both the past cases and future cases instruments use more than 10 observations to estimate leniency. The outcomes are annual labor income in SEK for 2019 prices and having any employment in a given year averaged over the seven years following the start of the debt restructuring program. I control for unit times year by using fixed effects in the OLS regression and the 2SLS estimation using the leave-one-out instruments and by using a leave-examiner-out residualization in the 2SLS estimation using past-cases-only and future cases. Standard errors are clustered on the applicant level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Income	Income	Income	Income	Income	Income	Income	Income
Accepted	-14389.9	-225.8	-20604.0	-13031.6	-21918.1	-25134.6	-19169.1	-14389.9
	(13310.0)	(12259.9)	(12223.8)	(14836.3)	(18185.0)	(22063.6)	(15666.2)	(15092.7)
Relative to control mean	-0.091	-0.002	-0.127	-0.083	-0.136	-0.154	-0.119	-0.091
Control for Year x Unit	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Balanced sample	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Min. number of cases	10	10	10	25	75	125	10	10
Past cases lag	30	30	30	30	30	30	60	30
Bootstrap SE	No	No	No	No	No	No	No	Yes
Observations	17032	21175	17032	15144	10424	7241	16312	17032

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13: Robustness - past-cases-only IV - income

Note: This table shows coefficients for 2SLS regressions for the instrument using the past-cases-only IV. The outcome is annual labor income in SEK for 2019 prices averaged over the seven years following the start of the debt restructuring program. Column 1 shows the baseline estimate. Column 2 drops the restriction on keeping a balanced sample. Column 3 drops the residualization by unit times year. Columns 4 to 6 increase the minimum number of cases from which the past-cases-only instrument is constructed to 25, 75, and 125, respectively. In column 6 cases are considered as past if the decision was made 60 days or more in the past. Standard errors are clustered on the applicant level.

	(1) Employed	(2) Employed	(3) Employed	(4) Employed	(5) Employed	(6) Employed	(7) Employed	(8) Employed
Accepted	-0.0917*	-0.0654*	-0.109**	-0.121**	-0.180***	-0.145*	-0.0965*	-0.0917*
	(0.0369)	(0.0331)	(0.0338)	(0.0418)	(0.0515)	(0.0634)	(0.0444)	(0.0429)
Relative to control mean	-0.140	-0.101	-0.165	-0.181	-0.255	-0.213	-0.147	-0.140
Control for Year x Unit	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Balanced sample	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Min. number of cases	10	10	10	25	75	125	10	10
Past cases lag	30	30	30	30	30	30	60	30
Bootstrap SE	No	No	No	No	No	No	No	Yes
Observations	17032	21243	17032	15144	10424	7241	16312	17032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A14: Robustness - past-cases-only IV - employment

Note: This table shows coefficients for 2SLS regressions for the instrument using the past-cases-only IV. The outcome is having any employment in a given year averaged over the seven years following the start of the debt restructuring program. Column 1 shows the baseline estimate. Column 2 drops the restriction on keeping a balanced sample. Column 3 drops the residualization by unit times year. Columns 4 to 6 increase the minimum number of cases from which the past-cases-only instrument is constructed to 25, 75, and 125, respectively. In column 6 cases are considered as past if the decision was made 60 days or more in the past. In column 6 cases are considered as past if the decision was made 60 days or more in the past. Standard errors in column 8 are computed via a Bayesian bootstrap that takes the construction of the instrument into account. Standard errors are clustered on the applicant level.

	(1) Income	(2) Income	(3) Income	(4) Income	(5) Income	(6) Income	(7) Income	(8) Income
Accepted	-25888.1*	-13024.4	-30096.4**	-20084.5	-24684.2	-37516.6*	-22666.3	-25888.1*
	(11095.6)	(10151.7)	(10095.3)	(12648.8)	(15596.7)	(18531.8)	(11781.8)	(12516.2)
Relative to control mean	-0.157	-0.085	-0.179	-0.125	-0.151	-0.219	-0.139	-0.157
Control for Year x Unit	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Balanced sample	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Min. number of cases	10	10	10	25	75	125	10	10
Past cases lag	30	30	30	30	30	30	60	30
Bootstrap SE	No	No	No	No	No	No	No	yes
Observations	17032	21175	17032	15144	10424	7241	16360	17032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A15: Robustness - leave-one-out IV - income

Note: This table shows coefficients for 2SLS regressions for the instrument using the leave-one-out IV. The outcome is annual labor income in SEK for 2019 prices averaged over the seven years following the start of the debt restructuring program. Column 1 shows the baseline estimate. Column 2 drops the restriction on keeping a balanced sample. Column 3 drops the residualization by unit times year. Columns 4 to 6 increase the minimum number of cases from which the past-cases-only instrument is constructed to 25, 75, and 125, respectively. In column 6 cases are considered as past if the decision was made 60 days or more in the past. Standard errors in column 8 are computed via a Bayesian bootstrap that takes the construction of the instrument into account. Standard errors are clustered on the applicant level.

	(1) Employed	(2) Employed	(3) Employed	(4) Employed	(5) Employed	(6) Employed	(7) Employed	(8) Employed
Accepted	-0.147***	-0.117***	-0.148***	-0.151***	-0.195***	-0.212***	-0.141***	-0.147***
	(0.0313)	(0.0278)	(0.0283)	(0.0359)	(0.0446)	(0.0536)	(0.0335)	(0.0356)
Relative to control mean	-0.214	-0.172	-0.215	-0.220	-0.273	-0.293	-0.207	-0.214
Control for Year x Unit	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Balanced sample	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Min. number of cases	10	10	10	25	75	125	10	10
Past cases lag	30	30	30	30	30	30	60	30
Bootstrap SE	No	No	No	No	No	No	No	Yes
Observations	17032	21243	17032	15144	10424	7241	16360	17032

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A16: Robustness - leave-one-out IV - employment

Note: This table shows coefficients for 2SLS regressions for the instrument using the leave-one-out IV. The outcome is having any employment in a given year averaged over the seven years following the start of the debt restructuring program. Column 1 shows the baseline estimate. Column 2 drops the restriction on keeping a balanced sample. Column 3 drops the residualization by unit times year. Columns 4 to 6 increase the minimum number of cases from which the past-cases-only instrument is constructed to 25, 75, and 125, respectively. In column 6 cases are considered as past if the decision was made 60 days or more in the past. Standard errors in column 8 are computed via a Bayesian bootstrap that takes the construction of the instrument into account. Standard errors are clustered on the applicant level.

	High income		Low income	
	(1) During program	(2) After program	(3) During program	(4) After program
Accepted	-0.00915	0.212***	-0.0489	0.0256
	(0.0339)	(0.0470)	(0.0440)	(0.0615)
Constant	0.185***	0.129***	0.137***	0.111*
	(0.0194)	(0.0265)	(0.0316)	(0.0437)
Relative to control mean	-0.049	1.646	-0.357	0.230
Control for Year x Unit	Yes	Yes	Yes	Yes
Observations	5402	5402	4716	4716

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A17: Car purchases during and after the program

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are dummies for having purchased a car in a given year during the five years after the program began and for the two years after the program ended. I split the sample into tertiles using the average income in the two years prior to applications residualized by calendar year and report results for the highest and lowest group. I control for unit times year by using a leave-examiner-out residualization. Standard errors are clustered on the applicant level.

	Low predicted garnishment		High predicted garnishment	
	(1)	(2)	(3)	(4)
	Has garnishment	Income	Has garnishment	Income
Accepted	-0.120	-30282.6	-0.0919	-42526.0
	(0.0642)	(29277.6)	(0.0534)	(23377.9)
Constant	0.159**	57743.8**	0.153***	69582.2***
	(0.0487)	(22102.3)	(0.0360)	(15698.3)
Relative to control mean	-0.758	-0.524	-0.602	-0.611
Control for Year x Unit	Yes	Yes	Yes	Yes
Observations	2336	2336	2333	2333

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A18: By predicted wage garnishment - low income

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are the average probability of being subject to wage garnishment in the seven years after the program started and average income measured in SEK in 2019 prices for low-income applicants. I split the sample and use the first sub-sample to predict the probability of being subject to wage garnishment out-of-sample in the second sub-sample and vice versa. To predict wage garnishment, I estimate an OLS regression of having wage garnishment on 5 bins of pre-application income and years of education, 10 bins of age in the year before application and dummies for gender, having a spouse, and having children as well as the year of application. Columns 1 and 3 show the actual share of individuals under wage garnishment for applicants with low and high predicted wage garnishment respectively, and columns 2 and 4 the effects on income measured in SEK in 2019 prices. Standard errors are clustered on the applicant level.

	Hours per week			Different firm		
	(1)	(2)	(3)	(4)	(5)	(6)
	High income	Medium income	Low income	High income	Medium income	Low income
Accepted	3.300*	-4.286	-0.971	0.00152	-0.0950	-0.229
	(1.597)	(2.831)	(5.653)	(0.0555)	(0.0715)	(0.261)
Constant	29.45***	28.33***	28.62***	0.340***	0.632***	0.720***
	(0.923)	(1.696)	(3.504)	(0.0316)	(0.0412)	(0.133)
Relative to control mean	0.112	-0.151	-0.034	0.004	-0.150	-0.318
Control for Year x Unit	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4004	2390	813	5391	3474	235

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A19: Intensive margin labor supply

Note: This table shows coefficients for 2SLS regressions using the past-cases-only instrument. The outcomes are weekly hours worked and the average of a dummy for working in a different firm as in the year before application for the seven years following the start of the debt restructuring program. Hours are only observed if the applicant works and if they appear in the wage structure survey. Working for the same firm is only defined until 2019 and if the applicant as employed in the year before application. I split the sample into tertiles using the average income in the two years prior to applications residualized by calendar year. I control for unit times year by using a leave-examiner-out residualization in the 2SLS estimation using past-cases-only. Standard errors are clustered on the applicant level.

A2 Figures

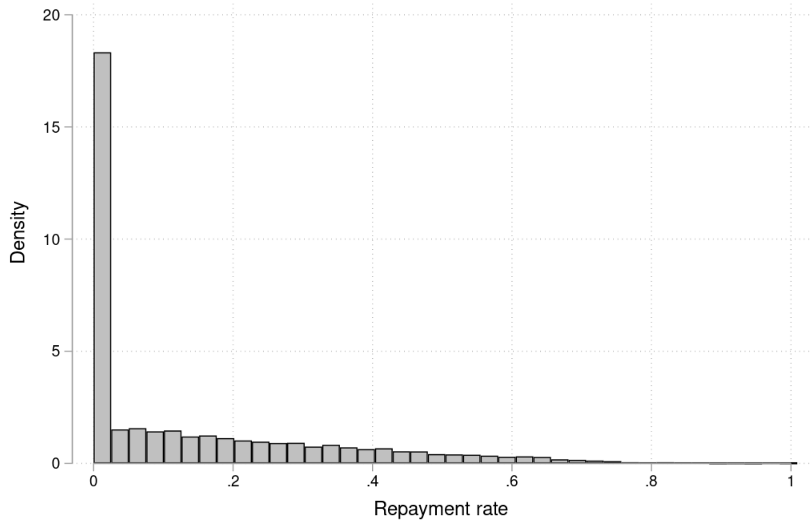
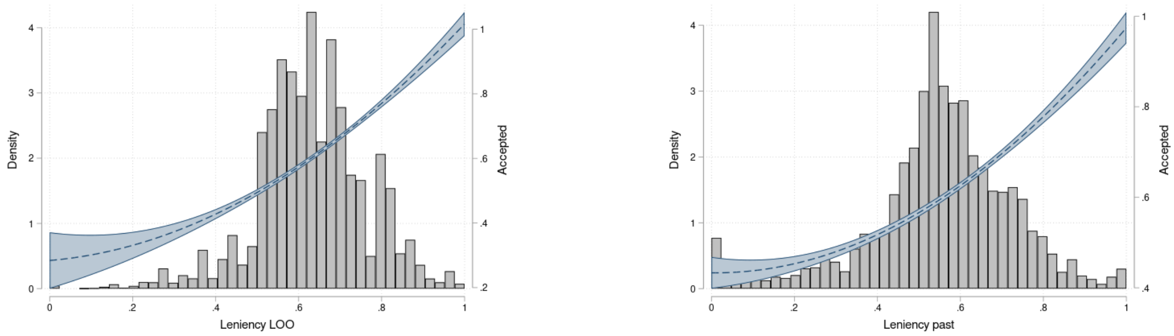


Figure A10: Repayment rate of accepted applicants

Note: This figures shows histogram of accepted applicants repayment rates according to their initial repayment plan. The repayment rate is defined as the overall amount of debt that must be repaid over the entire program divided by the total amount of debt that will be discharged after the program.

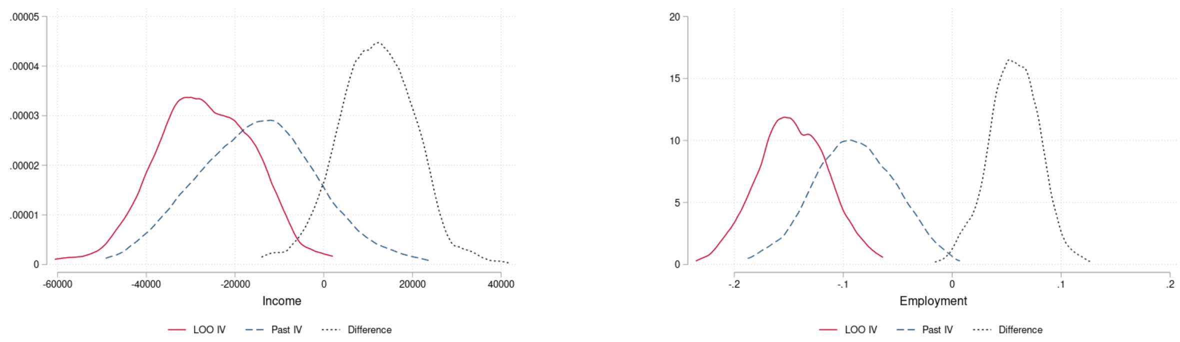


(a) leave-one-out leniency

(b) Past leniency

Figure A11: Distribution of leniency instruments

Note: These figures show histograms of the leniency instruments based on leave-one-out averages (a) and past averages (b) residualized by year times office using a leave-examiner-out approach. The left y-axes show a smoothed fit of the first stage displaying the average acceptance rate for each level of leniency.



(a) Income

(b) Employment

Figure A12: Bootstrap distributions of 2SLS estimations

Note: These figures show bootstrap distributions of the 2SLS estimates using the leave-one-out IV and the past-cases-only IV, as well as the difference between the two instruments ($\hat{\beta}_{\text{LOO}} - \hat{\beta}_{\text{past}}$). The bootstrap uses the exchangeably weighted bootstrap (Praestgaard and Wellner, 1993) clustered on the applicant level with 500 repetitions each.

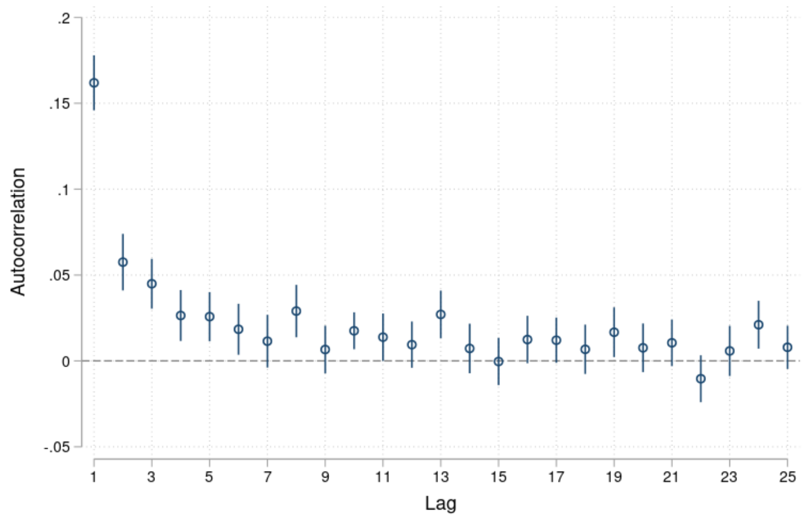
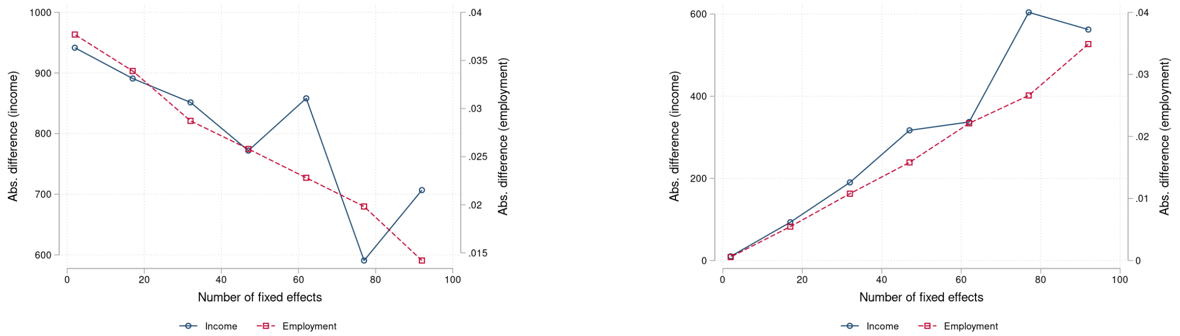


Figure A13: Autocorrelation of examiner decisions

Note: This figures shows the coefficients from an AR(25) model on examiners decisions using the bias-corrected method of moments estimator from Breitung, Kripfganz, and Hayakawa (2022). Standard errors are robust.



(a) $|\hat{\beta}_{\text{LOO}} - \bar{\beta}_{\text{past}}|$

(b) $|\hat{\beta}_{\text{past}} - \bar{\beta}_{\text{past}}|$

Figure A14: Simulate many controls bias

Note: These figures show the average absolute distances between the leave-one-out estimator with group fixed effects and the past-cases-only estimator using a leave-examiner-out residualization $\hat{\beta}_{\text{LOO}} - \bar{\beta}_{\text{past}}$ and the past-cases-only estimator using fixed effects and the residualized past-cases-only estimator $\hat{\beta}_{\text{past}} - \bar{\beta}_{\text{past}}$. For each number of fixed effects, I randomly assign examiners into fixed effects groups 500 times and report the average estimate for those draws. The outcomes are average income over the next seven years in SEK in 2019 prices and average probability of having any employment over the next seven years after the debt restructuring program begins.

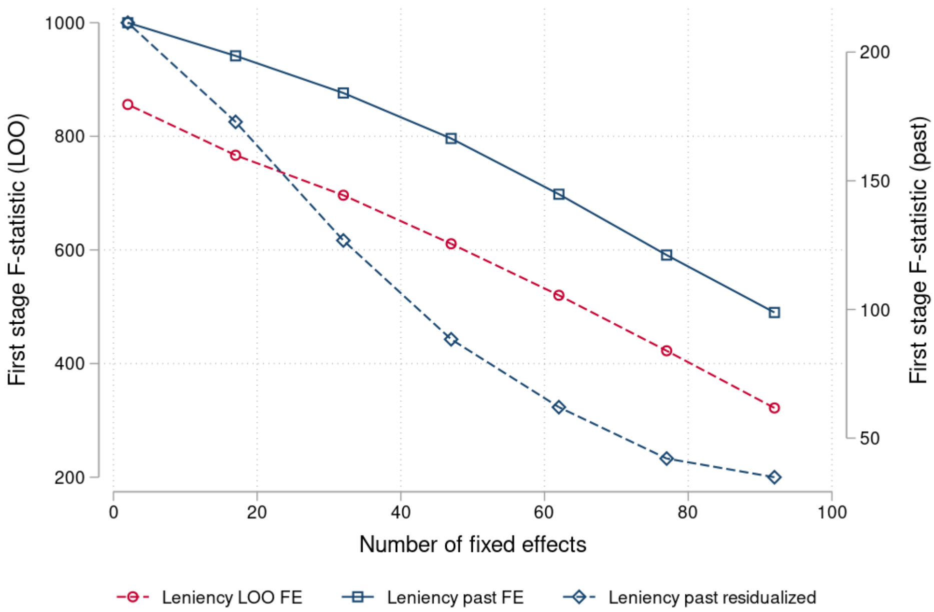


Figure A15: First stage F-statistics in many controls simulation

Note: This figures shows the average first stage F-statistics for the regressions of the examiner's decision of (i) the leave-one-out leniency instrument conditional on random group fixed effects (ii) the past-cases-only leniency IV conditional on random group fixed effects and (iii) the past-cases-only leniency IV residualized leaving the respective examiner out within random fixed effect group. For each number of fixed effects, I randomly assign examiners into fixed effects groups 500 times and report the average F-statistic for those draws.

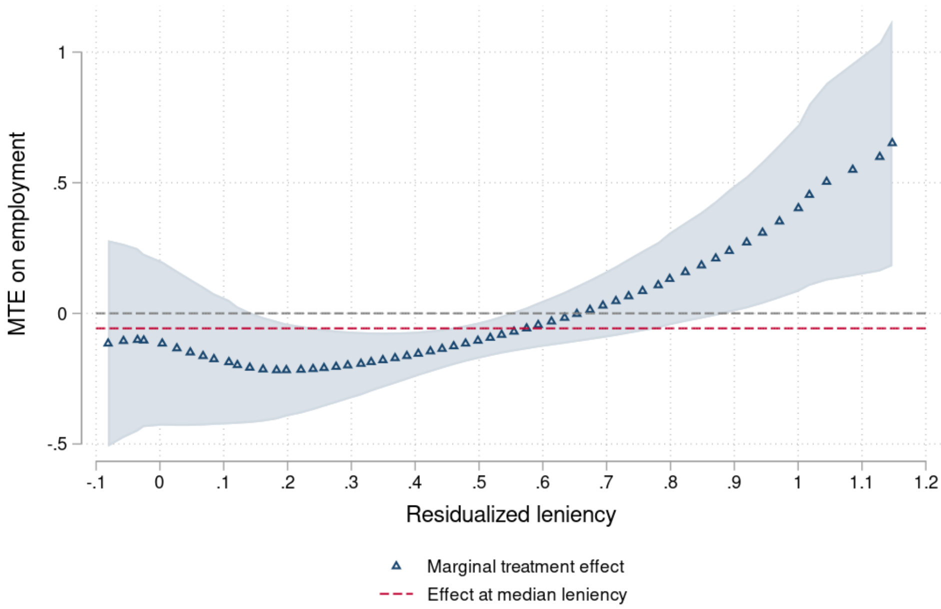
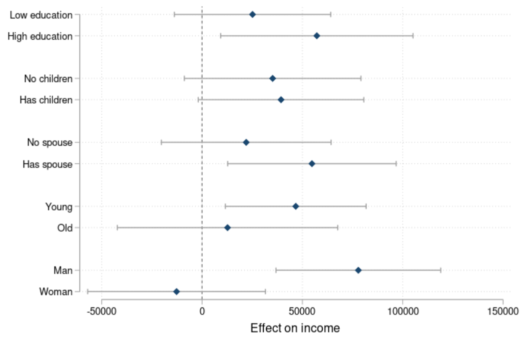
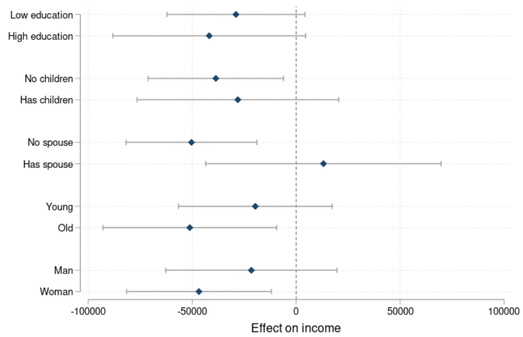


Figure A16: Marginal treatment effects by estimated leniency

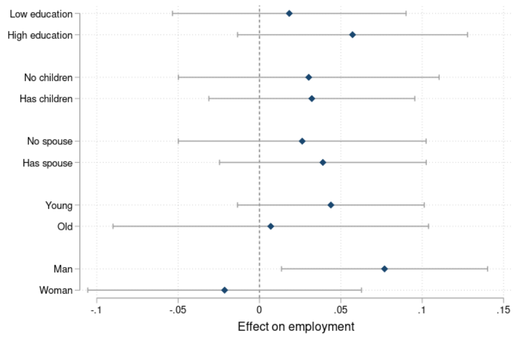
Note: This figures plots the distribution of marginal treatment effects estimated on average employment in the seven years after the program starts using a semi-parametric local IV estimation (Andresen, 2018) against estimated residualized leniency using past-cases-only. Leniency is residualized by year times office using a leave-examiner-out approach adding average leniency. The dotted red line marks the marginal treatment effect at the median leniency in the sample. 95 percent confidence intervals are computed by bootstrapping with 300 repetitions.



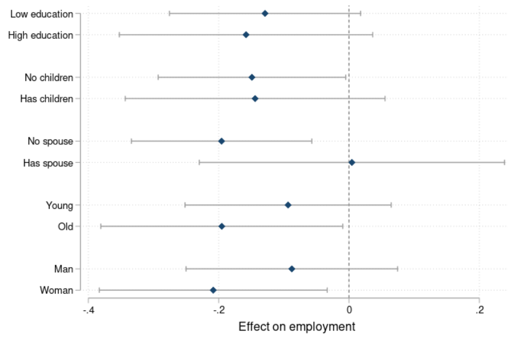
(a) Income - high income applicants



(b) Income - low income applicants



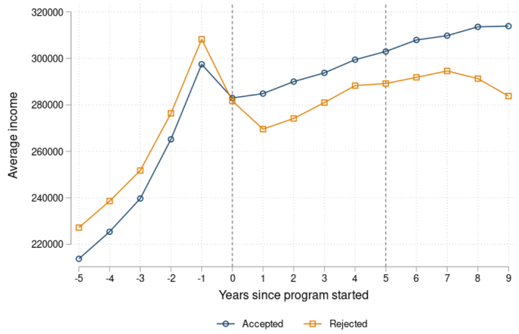
(c) Employment - high income applicants



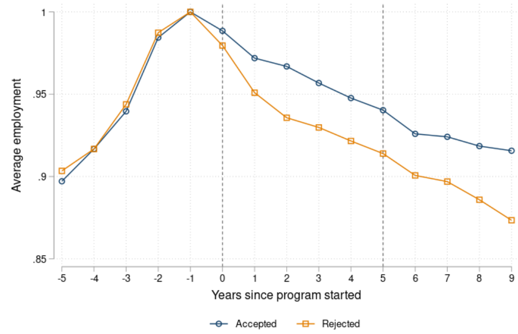
(d) Employment - low income applicants

Figure A17: Average income and employment- heterogeneity by income

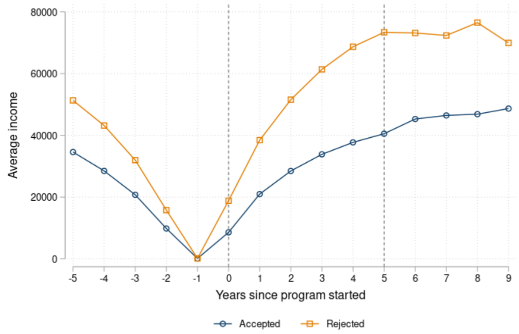
Note: These figures show the coefficients and 95 percent confidence intervals from estimating the 2SLS regression using the past-cases-only instrument. The sample is split twice. Panels (a) and (c) sow results for above-median initial income applicants and panels (b) and (d) for below-median initial income applicants. Within these splits, I further split the sample by applicant characteristics measured in the year prior to application. The first split is by below and above median years of education, the second split is by having any children, the third split is by cohabiting with a spouse. The fourth split is by below and above median age, and the fifth by gender. The outcomes are average annual labor income in SEK for 2019 prices and the average of a dummy for having any employment in a given year for the seven years following the start of the debt restructuring program. Standard errors are clustered on the applicant level.



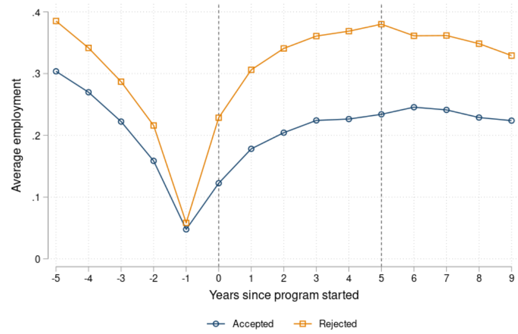
(a) Income - high income applicants



(b) Employment - high income applicants



(c) Income - low income applicants



(d) Employment low income applicants

Figure A18: Average income and employment- heterogeneity by income

Note: These figures show average income and employment for accepted and rejected applicants splitting the sample by initial income. For that I estimate average income in the year prior to application residualized by calendar year. I then split the sample into tertiles. Panels (a) and (b) show results for the highest income tertile and panels (c) and (d) for the lowest income tertile. The outcomes are annual labor income in SEK in 2019 prices and an indicator for having had any employment in the given year.

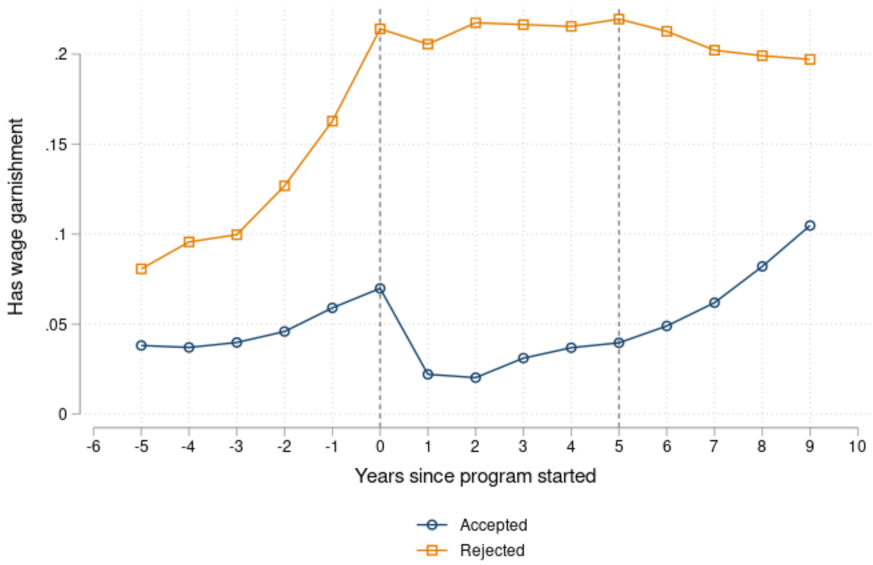


Figure A19: Share of applicants under wage garnishment

Note: This figures plots share of accepted and rejected applicants who are subject to wage garnishment relative to the year of the decision.

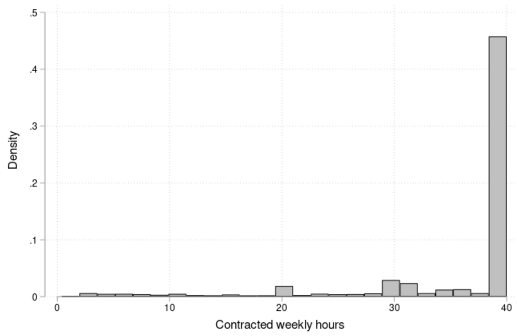


Figure A20: Distribution of contracted hours

Note: This figures shows histogram of contracted weekly hours as reported in the Structure of Earnings Survey

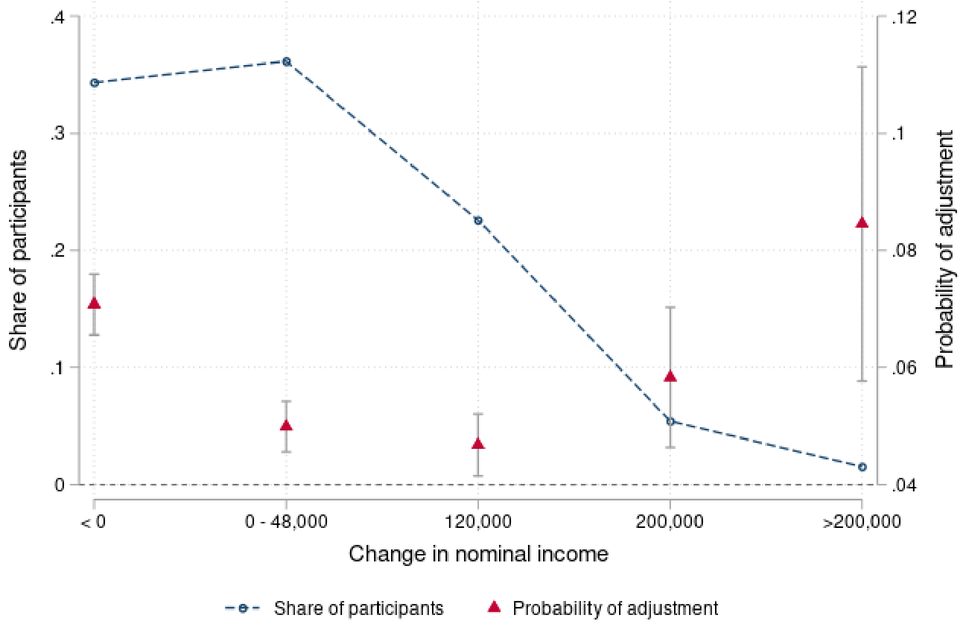


Figure A21: Adjustment by change in income for initially employed

Note: The red triangles show average rates of adjustments of the repayment plan in the same or following year for bins of changes in nominal income for participants with initial labor income above 160,000 SEK. (right y-axis). The first bin includes everyone who experiences a drop in income, the second bin includes participants whose income does not change or changes by less than 48,000 SEK, the third bin participants whose income increase between 48,000 and 120,000 SEK, the fourth bin participants whose income increase between 120,000 and 200,000 SEK, and the last bin all participants with income increases of more than 200,000 SEK. The blue dotted line shows the share of participants in the respective bin.

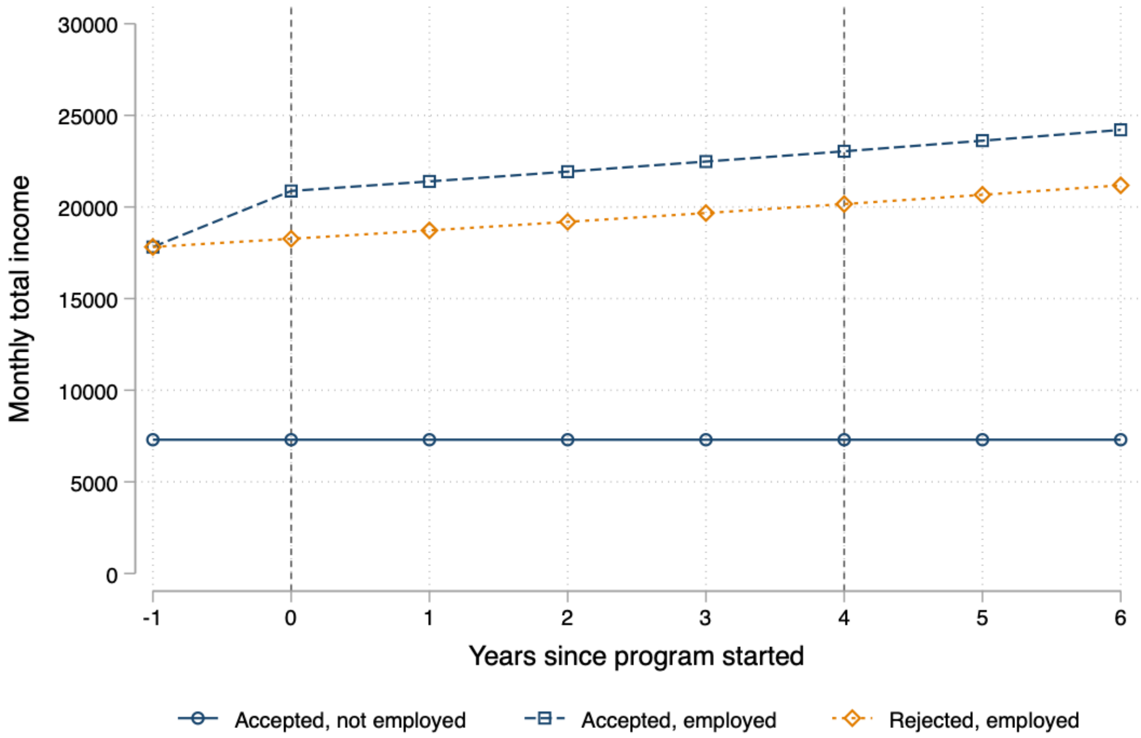


Figure A22: Monthly income - model

Note: The graphs show monthly hours predicted by the model. Income includes both labor income and social benefits. Initial hours in period -1 are set endogenously. The dotted orange line plots outcomes for a rejected applicant who initially works 35 hours per week and who faces wage garnishment. The dashed blue line plots outcomes for an accepted applicant who initially works 35 hours per week. And the solid blue line plots outcomes for an accepted applicant who does not work in the year prior to application. The first vertical dashed line indicated the begin of the debt restructuring program and the second dashed line its end.

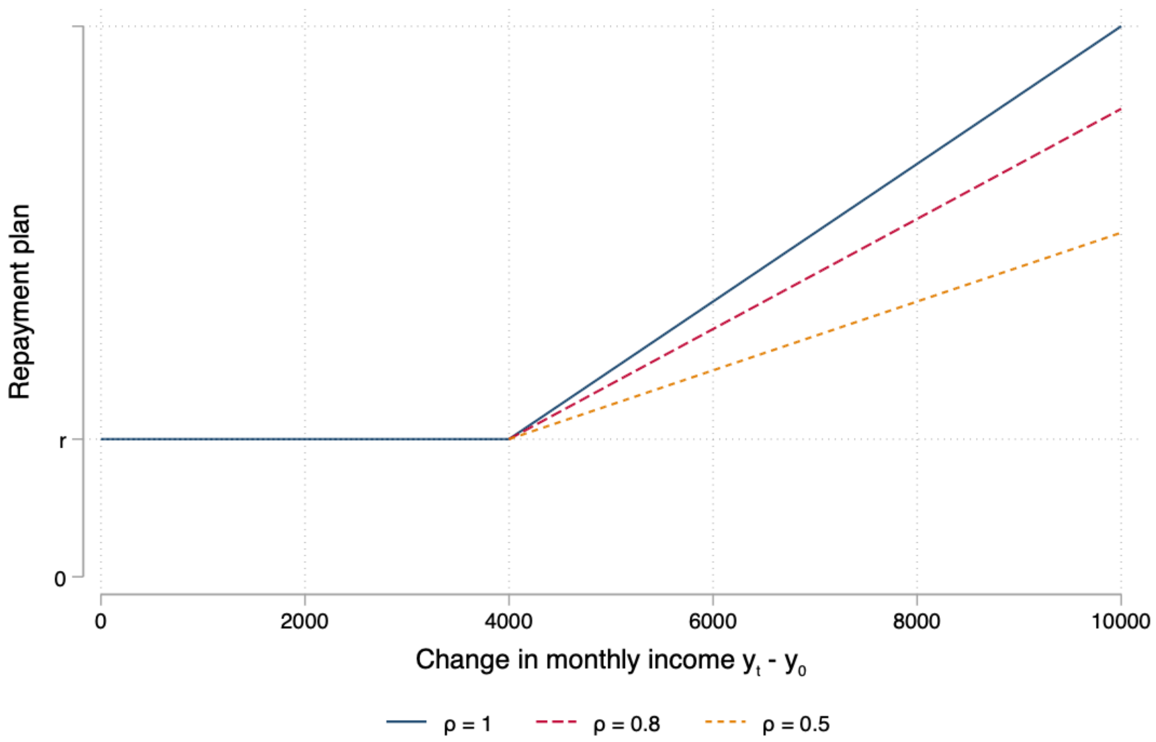


Figure A23: Repayment for changes in income

Note: The graphs plots the repayment amount for different changes in income relative to the reference income and different marginal adjustment rates ρ as in equation (20).

B1 Examiner IV - extensions and derivations

B1.1 Asymptotic bias with many controls

Many empirical applications require the inclusion of some controls for identification. In the following section I will show how, similar to Kolesár (2013), this can reintroduce the correlation between unobserved case characteristics U_{jt} and future decisions D_{jk} even when using only past cases as instrument. Kolesár (2013) shows that including many controls can lead to inconsistent estimates in standard jackknife IV estimations.

Let $W_{jt} \in \mathcal{W}$ be some arbitrary discrete group to which we assign observation (j,t). For simplicity, I assume that W_{jt} is orthogonal to all decisions and therefore irrelevant from an identification point of view

$$E[W_{ik}D_{jt}] = 0 \text{ for all } k \in \mathcal{T} \text{ and all } i \in \mathcal{J} \ . \tag{A.6}$$

In empirical applications W_{jt} could, for example, refer to court-by-year fixed effects. Let N_W be the number of groups, N_J the number of examiners in each group, and N_T the number of cases for a given group and examiner. To simplify notation I assume that N_J and N_T are equal across all groups and examiners. Let $\tilde{Z}_{jt} = Z_{jt} - \frac{1}{N_J \times N_T}\sum_i \sum_s Z_{is}$ denote the instrument de-meaned by group fixed effects.[46] The past IV estimator can then be written as

$$\tilde{\beta}_{\text{past}} = \frac{\sum_j \sum_t \tilde{Z}_{jt}^{\text{past}} Y_{jt}}{\sum_j \sum_t \tilde{Z}_{jt}^{\text{past}} D_{jt}} \tag{23}$$

To derive the asymptotic bias, I follow Kolesár (2013) and let only the number of groups N_W grow to infinity while keeping the number of examiners within each group N_W and the number of cases within each examiner-group N_T fixed. Then as $N_W \to \infty$

$$\tilde{\beta}_{\text{past}} \xrightarrow{p} \beta - \frac{\alpha}{N_J \times N_T} \frac{\sum_t \sum_s \frac{1}{s-1} \sum_{l \in \mathcal{P}_s} E\left[D_{jl} U_{jt}\right]}{\sum_t E\left[\tilde{Z}_{jt}^{\text{past}} D_{jt}\right]} \tag{24}$$

By averaging over all other instruments within a fixed effects group in the residualization step, we may capture other observations whose instrument is constructed from current or future decisions of the same examiner. These decisions are exactly those that we want to omit because of their correlation with the regression error. A simple solution to this issue is to use a leave-examiner-out residualization instead $\bar{Z}_{jt} = Z_{jt} - \frac{1}{(N_J-1)\times N_T}\sum_{i\neq j}\sum_s Z_{is}$. This ensures that the instrument is residualized using only observations exogenous to the respective examiner's

[46] For brevity the index i in the first sum refers to all examiners within the same groups such that $W_{ik} = W_{jt}$ for at least one k, including examiner j. The second sum sums over all N_T cases s for which a given examiner falls into this group.

decisions. Then, the corresponding estimator will be able to recover the true causal effect

$$\bar{\beta}_{\text{past}} = \frac{\sum_j \sum_t \bar{Z}^{\text{past}}_{jt} Y_{jt}}{\sum_j \sum_t \bar{Z}^{\text{past}}_{jt} D_{jt}} \xrightarrow{p} \beta. \tag{25}$$

B1.2 Heterogeneous treatment effects

It is a well-known result that, with treatment effect heterogeneity, 2SLS regressions do not necessarily identify average treatment effects of the underlying population, but weighted averages of compliers (Imbens and Angrist, 1994). Differences in estimates between the leave-one-out IV and the past-cases-only IV can in such settings therefore arise both due of biases in the estimate as well as different weights attached to each observation. To model treatment effect heterogeneity, I allow β to be a function of unobserved case characteristics and impose the assumption that conditional on case characteristics, treatment effects are independent from all decisions and therefore the instruments.

$$\beta(U_{jt}) = E[\beta \mid U_{jt}] \text{ such that } \beta(U_{jt}) \perp\!\!\!\perp D_{jk} \mid U_{jt} \text{ for all } k \in \mathcal{T} . \tag{A.4}$$

The respective probability limits of the estimators are given by

$$\hat{\beta}_{\text{LOO}} \xrightarrow{p} E[\omega_l(U_{jt})\beta(U_{jt})] + \alpha \frac{\sum_t \sum_{k \in \mathcal{P}_t} E\left[D_{jk} U_{jt}\right]}{\sum_t \sum_{k \neq t} E\left[D_{jk} D_{jt}\right]} \tag{26}$$

and

$$\hat{\beta}_{\text{past}} \xrightarrow{p} E[\omega_p(U_{jt})\beta(U_{jt})]. \tag{27}$$

If we let U_{jt} follow a discrete support $\mathcal{U}$ the weights are

$$\omega_l(u) = \frac{\sum_t \frac{1}{T-1} \sum_{k \neq t} Pr(U_{jt} = u) E[D_{jk} D_{jt} \mid U_{jt} = u]}{\sum_{w \in \mathcal{U}} \sum_t \frac{1}{T-1} \sum_{k \neq t} Pr(U_{jt} = w) E[D_{jk} D_{jt} \mid U_{jt} = w]} \tag{28}$$

and

$$\omega_p(u) = \frac{\sum_t \frac{1}{t-1} \sum_{k < t} Pr(U_{jt} = u) E[D_{jk} D_{jt} \mid U_{jt} = u]}{\sum_{w \in \mathcal{U}} \sum_t \frac{1}{t-1} \sum_{k < t} Pr(U_{jt} = w) E[D_{jk} D_{jt} \mid U_{jt} = w]}. \tag{29}$$

The main difference is that the weight in the leave-one-out estimation sums over all past and future decisions while the weight in the past-cases-only regressions naturally only sums over past cases. Systematic differences between the two weights therefore depend on how the conditional expectation $E[D_{jt} D_{jk} \mid U_{jt}]$ evolves over time within examiner. The average decision used in ω_l will by construction capture later cases than the average decision in ω_p. In an examiner decision model such as (2) where both examiner leniency and the mapping between leniency and case characteristics to decisions are constant over time, both weights will yield identical results. If, for example, examiner leniency grows over time such that $E[D_{jt}D_{jt-p} \mid U_{jt}] \neq E[D_{jt}D_{jt+p} \mid U_{jt}]$ and these differences are systematically related to $\beta(U_{jt})$, then comparisons between the two estimators can capture both the bias and differences in weighting.

A minimal requirement that is often stated for weighted treatment effects to be interpreted causally is that the estimate should represent a convex combination of underlying treatment effects (see e.g. de Chaisemartain and D'Haultfœuille, 2020). To ensure non-negative weights for every observation, we can impose the following monotonicity assumption

$$E[Z_{jt}^{\text{past}} D_{jt} \mid U_{jt}] \geq 0. \tag{A.5}$$

This condition implies that, for any level of characteristics U_{jt}, uptake of the program increases in leniency. The regression weights will then be zero for all always-takers and never-takers for who treatment status does not vary with leniency. Compliers will be weighted proportional to how strongly their treatment status reacts to changes in leniency of assigned examiners.

B1.3 Derivation of of technical results

B1.3.1 Bias leave-one-out IV without controls

The standard leave-one-out 2SLS estimator is given by

$$\hat{\beta}_{\text{LOO}} = \frac{\sum_j \sum_t Z_{jt}^{\text{LOO}} Y_{jt}}{\sum_j \sum_t Z_{jt}^{\text{LOO}} D_{jt}} = \frac{\sum_j \sum_t \frac{1}{T-1} \sum_{k \neq t} D_{jk} Y_{jt}}{\sum_j \sum_t \frac{1}{T-1} \sum_{k \neq t} D_{jk} D_{jt}} \tag{30}$$

where the second step follows from the definition of the leave-one-out instrument (8). Then when we let the number of examiner s grow to infinity $J \to \infty$

$$\begin{aligned}\hat{\beta}_{\text{LOO}} \xrightarrow{p} \frac{E\left[\sum_t \frac{1}{T-1} \sum_{k \neq t} D_{jk} Y_{jt}\right]}{E\left[\sum_t \frac{1}{T-1} \sum_{k \neq t} D_{jk} D_{jt}\right]} &= \frac{E\left[\sum_t \sum_{k \neq t} D_{jk} Y_{jt}\right]}{E\left[\sum_t \sum_{k \neq t} D_{jk} D_{jt}\right]} \\ = \beta + \alpha \frac{\sum_t \sum_{k \neq t} E\left[D_{jk} U_{jt}\right]}{\sum_t \sum_{k \neq t} E\left[D_{jk} D_{jt}\right]} &= \beta + \alpha \frac{\sum_t \sum_{k \in \mathcal{P}_t} E\left[D_{jk} U_{jt}\right]}{\sum_t \sum_{k \neq t} E\left[D_{jk} D_{jt}\right]}\end{aligned} \tag{31}$$

The third step follows from the structural model (3) and assumption (A.1) the last step from assumption (A.2).

B1.3.2 Regression weights with heterogeneous treatment effects

If we allow β to vary with case characteristics U_{jt}, then similar to (31)

$$
\begin{aligned}
\hat{\beta}_{\text{LOO}} &\xrightarrow{p} \frac{E\left[\sum_t \frac{1}{T-1}\sum_{k\neq t} D_{jk}Y_{jt}\right]}{E\left[\sum_t \frac{1}{T-1}\sum_{k\neq t} D_{jk}D_{jt}\right]} \\
&= \frac{E\left[\sum_t \sum_{k\neq t} D_{jk}Y_{jt}\right]}{E\left[\sum_t \sum_{k\neq t} D_{jk}D_{jt}\right]} \\
&= \frac{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\beta\right]}{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\right]} + \alpha\frac{\sum_t \sum_{k\neq t} E\left[D_{jk}U_{jt}\right]}{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\right]} \\
&= \frac{\sum_t \sum_{k\neq t} E\left[E\left[D_{jk}D_{jt}\beta \mid U_{jt}\right]\right]}{\sum_t \sum_{k\neq t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\right]} + \alpha\frac{\sum_t \sum_{k\in\mathcal{P}_t} E\left[D_{jk}U_{jt}\right]}{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\right]} \\
&= \frac{\sum_t \sum_{k\neq t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\beta(U_{jt})\right]}{\sum_t \sum_{k\neq t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\right]} + \alpha\frac{\sum_t \sum_{k\in\mathcal{P}_t} E\left[D_{jk}U_{jt}\right]}{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\right]} \\
&= E\left[\frac{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt} \mid U_{jt}\right]\beta(U_{jt})}{\sum_t \sum_{k\neq t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\right]}\right] + \alpha\frac{\sum_t \sum_{k\in\mathcal{P}_t} E\left[D_{jk}U_{jt}\right]}{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\right]} \\
&= \sum_{u\in\mathcal{U}} \frac{\sum_t \sum_{k\neq t} Pr(U_{jt}=u)E\left[D_{jk}D_{jt} \mid U_{jt}=u\right]}{\sum_{w\in\mathcal{U}}\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt} \mid U_{jt}=w\right]}\beta(u) + \alpha\frac{\sum_t \sum_{k\in\mathcal{P}_t} E\left[D_{jk}U_{jt}\right]}{\sum_t \sum_{k\neq t} E\left[D_{jk}D_{jt}\right]}
\end{aligned} \tag{32}
$$

where the third step follows from the law of iterated expectations and the fourth step follows from assumption (A.4). We then assume that U_{jt} has discrete support $\mathcal{U}$ in the sixth step. Similarly, we can derive the weights for the past-cases-only estimator

$$
\begin{aligned}
\hat{\beta}_{\text{past}} &\xrightarrow{p} \frac{E\left[\sum_t \frac{1}{t-1}\sum_{k<t} D_{jk}Y_{jt}\right]}{E\left[\sum_t \frac{1}{t-1}\sum_{k<tt} D_{jk}D_{jt}\right]} \\
&= \frac{E\left[\sum_t \frac{1}{t-1}\sum_{k<t} D_{jk}Y_{jt}\right]}{E\left[\sum_t \frac{1}{t-1}\frac{1}{t-1}\sum_{k<t} D_{jk}D_{jt}\right]} \\
&= \frac{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[D_{jk}D_{jt}\beta\right]}{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[D_{jk}D_{jt}\right]} \\
&= \frac{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[E\left[D_{jk}D_{jt}\beta \mid U_{jt}\right]\right]}{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\right]} \\
&= \frac{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\beta(U_{jt})\right]}{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\right]} \\
&= E\left[\frac{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[D_{jk}D_{jt} \mid U_{jt}\right]\beta(U_{jt})}{\sum_t \frac{1}{t-1}\sum_{k<t} E\left[E\left[D_{jk}D_{jt} \mid U_{jt}\right]\right]}\right] \\
&= \sum_{u\in\mathcal{U}} \frac{\sum_t \frac{1}{t-1}\sum_{k<t} Pr(U_{jt}=u)E\left[D_{jk}D_{jt} \mid U_{jt}=u\right]}{\sum_{w\in\mathcal{U}}\sum_t \frac{1}{t-1}\sum_{k<t} E\left[D_{jk}D_{jt} \mid U_{jt}=w\right]}\beta(u)
\end{aligned} \tag{33}
$$

B1.3.3 Bias past IV with controls

When we use the instrument derived from past-cases-only (11) with fixed effects W, then the estimator can be written as

$$\tilde{\beta}_{\text{past}} = \frac{\sum_j \sum_t \tilde{Z}_{jt}^{\text{past}} Y_{jt}}{\sum_j \sum_t \tilde{Z}_{jt}^{\text{past}} D_{jt}} \tag{34}$$
$$= \frac{\sum_j \sum_t \left(\frac{1}{t-1} \sum_{k<t} D_{jk} - \frac{1}{N_J \times N_T} \sum_i \sum_s \frac{1}{s-1} \sum_{l<s} D_{il} \right) Y_{jt}}{\sum_j \sum_t \tilde{Z}_{jt}^{\text{past}} D_{jt}}.$$

Then as the number of controls increases keeping the number of examiners within fixed effects group and the number of cases within examiner - fixed effect group fixed we can derive the bias.

$$\tilde{\beta}_{\text{past}} \xrightarrow{p} \beta + \alpha \frac{E\left[\sum_t \left(\frac{1}{t-1} \sum_{k<t} D_{jk} - \frac{1}{N_J \times N_T} \sum_i \sum_s \frac{1}{s-1} \sum_{l<s} D_{il} \right) U_{jt}\right]}{E\left[\sum_t \tilde{Z}_{jt}^{\text{past}} D_{jt}\right]} \tag{35}$$
$$= \beta + \alpha \frac{\sum_t \left[\sum_{k<t} E\left[D_{jk} U_{jt}\right] - \frac{1}{N_J \times N_T} \left(\sum_{i \neq j} \sum_s \frac{1}{s-1} \sum_{l<s} E\left[D_{il} U_{jt}\right] + \sum_s \frac{1}{s-1} \sum_{l<s} E\left[D_{jl} U_{jt}\right] \right)\right]}{\sum_t E\left[\tilde{Z}_{jt}^{\text{past}} D_{jt}\right]}$$
$$= \beta - \frac{\alpha}{N_J \times N_T} \frac{\sum_t \sum_s \frac{1}{s-1} \sum_{l \in \mathcal{P}_s} E\left[D_{jl} U_{jt}\right]}{\sum_t E\left[\tilde{Z}_{jt}^{\text{past}} D_{jt}\right]}$$

where the last step follows from assumptions (A.2) and (A.6).

B1.3.4 Asymptotic bias with exogenous controls for leave-one-out estimator

Similarly to (35) we can derive the bias in the classical leave-one-out 2SLS estimator when using controls as above

$$\bar{\beta}_{\text{LOO}} \xrightarrow{p} \beta + \alpha \frac{E\left[\sum_t \left(\frac{1}{T-1} \sum_{k \neq t} D_{jk} - \frac{1}{N_J \times N_T} \sum_i \sum_s \frac{1}{T-1} \sum_{l \neq s} D_{il} \right) U_{jt}\right]}{E\left[\sum_t \tilde{Z}_{jt}^{\text{LOO}} D_{jt}\right]} \tag{36}$$
$$= \beta + \alpha \frac{E\left[\sum_t \left(\frac{1}{T-1} \sum_{k \in \mathcal{P}_t} D_{jk} - \frac{1}{N_J \times N_T} \sum_s \frac{1}{T-1} \sum_{l \in \mathcal{P}_s} D_{il} \right) U_{jt}\right]}{E\left[\sum_t \tilde{Z}_{jt}^{\text{LOO}} D_{jt}\right]}.$$

B2 Monte Carlo simulation

In this section, I simulate an examiner decision model and evaluate how the bias of the 2SLS estimation using the leave-one-out leniency instrument Z_{jt}^{LOO} varies with different parameters of the underlying model. I assume that examiners make a binary decision based on the case characteristics U_{jt}, the examiner's leniency, L_{jt}, and six lags of past case characteristics.

$$D_{jt} = \mathbf{1}\{\theta U_{jt} + \lambda L_j + \gamma \sum_{k=1}^{6} 0.5^k \times U_{jt-k} \geq 0\}. \tag{37}$$

The outcome is determined by the structural outcome model (3) such that

$$Y_{jt} = \beta D_{jt} + \alpha U_{jt} + \nu_{jt}. \tag{38}$$

All random variables L_j, U_{jt}, and ν_{jt} are iid $\mathcal{N}(0,1)$ distributed. I set the true causal parameter of interest β and the influence of case characteristics on the outcome α to 1. The initial decision parameters are given by $\theta = 0.8$, $\lambda = 0.1$, and $\gamma = 0.1$ to reflect that case characteristics are likely to play a substantially more important role in forming the decision than examiner's leniency or past cases. I set the number of examiners J to 5,000 and the number of cases per examiner T to 110. I then estimate the leave-one-out and past-cases-only instruments and restrict the sample to observations for which the instrument is estimated off at least 10 observations, that is, I drop the first 10 cases per examiner and keep 100 cases per examiner. Figure B1 shows the bias for both instruments, varying each one of the parameters $\{\theta,\ \lambda,\ \gamma,\ T\}$ while keeping the others fixed. Unsurprisingly, the bias induced by the leave-one-out instrument increases in the importance of lagged case characteristics γ and - given a fixed number of lags - decreases in the number of cases each examiner decides over. Perhaps surprisingly, however, the bias increases in the importance of current case characteristics θ and decreases in the importance of examiners' leniency λ. The reason for this is that the ratio of θ to λ governs the strength of the first stage. Since the bias is determined by the size of violations of the exclusion restriction scaled by the first stage, decreasing the latter while keeping the former fixed or letting it decrease at a slower rate, will exacerbate the bias in the leave-one-out estimation. In the past-cases-only estimation the exclusion restriction is not violated such that the size of the first stage does not matter for the bias in expectations, given that the relevance condition holds.

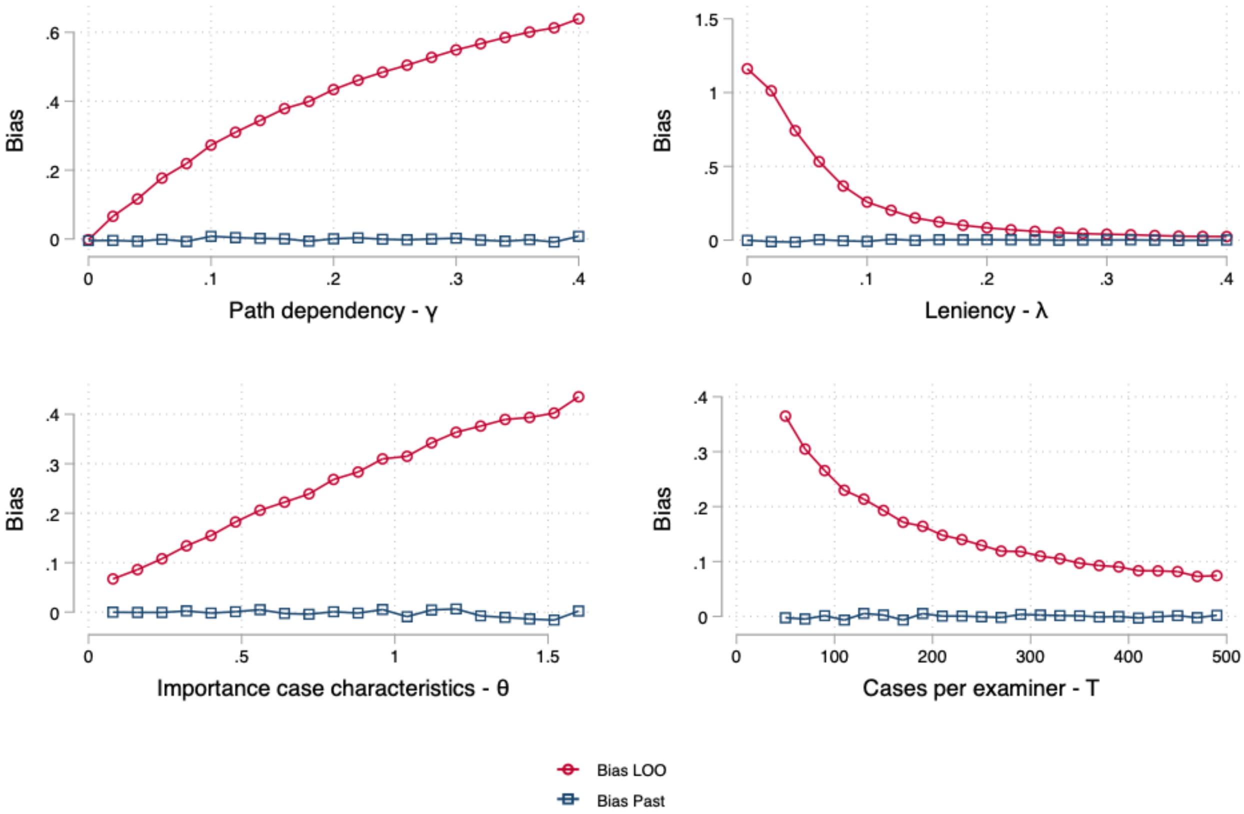


Figure B1: Results from Monte Carlo simulation

Note: This figure shows the bias for the leave-one-out 2SLS regression and the 2SLS regression using past-cases-only from on a Monte Carlo simulation based on the decision model (37). Each panel varies one of the three decision parameters in (37) or the number of cases per examiner and keeps the other three parameters fixed. Each point is the average estimate for 150 regressions.

C1 Bias in examiner IV - application to inventor mobility

A growing body of literature uses the quasi-random allocation of patent examiners to patent applications at the US Patent and Trademark Office (USPTO) to estimate the treatment effects of patent approvals (Galasso and Schankerman, 2015; Gaulé, 2018; Sampat and Williams, 2019; Melero et al., 2020). In the following section, I show that using the past-cases-only instrument instead of a leave-one-out instrument can substantially affect estimates of inventors' likelihood of remaining at their current firm. Melero et al. (2020) provide a theoretical framework in which gaining a monopoly through a patent approval increases firms' incentives to retain inventors. This, in turn, can limit inventor mobility and harm the diffusion of knowledge. Empirically, they instrument the number of patents an inventor holds with the average leave-one-out measure of patent examiner leniency over all patent applications of the given inventor. In this setting, an additional patent lowers the probability of changing employers by 2.3 percentage points. I use a specification with a binary indicator for patent approval as the endogenous variable which is identical to the specification in the econometric set-up in section 3. Using the leave-one-out instrument, I find that receiving a patent decreases the probability of working at the previous firm when applying for the next patent by 3 percentage points. This estimate more than doubles to a decrease of 6.6 percentage points when using the past-cases-only instrument. Relative to the OLS bias, this corresponds to a difference of 113 percent.

C1.1 Data

I download data on all patent applications from 2001 - 2019 from the USPTO.[47][48] These data contain information on the application date, if granted the date of filing, the respective examiner, the examiner's art unit group, and the applicant's name. I infer employer names from the assignment data keeping only those assignments that are marked as an "employer assign" by the USPTO. My main outcome is a dummy indicating that the inventor works at the same company in their next application as they do in their current application. This automatically drop all inventors with only one application. I harmonize employer names to some extent.[49] Nevertheless, the mobility measure is likely to overestimate mobility due to differences in company names' spelling e.g. because of typos or changes in their legal structure. This should not affect estimates if the measurement error in mobility is unrelated to examiner leniency. I follow the literature and assume that cases are randomly assigned to examiners within art unit. I therefore residualize the acceptance variable with the leave-examiner-out average acceptance

[47] Data on patent examiners are only available from 2001 onwards. I drop all applications after 2019 since the patent application process can take several years.

[48] https://www.uspto.gov/ip-policy/economic-research/research-datasets/patent-assignment-dataset

[49] E.g. I drop special characters or terms that refer to the legal structure such as "INC" from firm company names. These can contribute to small differences in spelling.

rate within the same art unit and year

$$D_{jt} = D_{jt} - \frac{1}{N_A(j) - N_j} \sum_{i:A(i)=A(j);\ i \neq j} D_{jt}. \tag{39}$$

The instruments are then constructed as averages over the residualized decisions. It can take several years until a patent is granted. This makes the issue of contemporaneity of cases more severe than in the debt restructuring setting. The average time between application and approval in my sample is 2.89 years. Around 99 percent of all patents are approved within 8 years. I therefore define the past-cases-only instrument as the average acceptance rate for all applications which were filed until the 8th calendar year before the respective application. I further ensure that all instruments are estimated using more than 10 cases. This leaves me with a sample of around 2.5 million patent applications and 8,884 examiners. The average acceptance rate is 84.9 percent. The average (median) leave-one-out instrument is estimated using 644 (488) other cases and the average (median) past cases instrument is estimated using 153 (182) cases.

C1.2 Results

Table C1 shows the OLS and 2SLS estimates for the two instruments. The OLS and 2SLS estimate based on the leave-one-out instrument are very similar, with -0.034 and -0.030 respectively. The past-cases-only estimator is substantially lower with -0.0657. Taking the past-cases-only estimator as ground truth, the bias of the leave-one-out estimator relative to the OLS bias corresponds to 113 percent.

	(1) OLS	(2) 2SLS LOO	(3) 2SLS past
Granted	-0.0342***	-0.0302*	-0.0657***
	(0.00123)	(0.0126)	(0.0184)
Constant	0.257***	0.253***	0.284***
	(0.00117)	(0.0108)	(0.0156)
Control Art Unit x Year	Yes	Yes	Yes
FS F-stat.		1398	632
Observations	2548909	2548909	2548909

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C1: OLS and 2SLS estimations

Note: This table shows coefficients for the OLS regressions and 2SLS regressions for each of the two instruments based on leave-one-out estimates of examiner leniency and past-cases-only. The outcome is the probability of working for the same employer at the time of the next patent application. I control for art unit times year by using fixed effects in the OLS regression and a leave-examiner out residualization of decision variables before aggregating them up to the instruments. I report the first stage F-statistics for the two 2SLS regressions. Standard errors are clustered on the examiner and applicant level.