Discussion of “Borrowers in the Shadows: The Promise and Pitfalls of Alternative Credit Data”

Benedict Guttman-Kenney

Rice University

WFA Conference, 23 June 2025

RICE | BUSINESS

What is alternative credit data ("Alt. data")?

All 3 credit reporting agencies have alt. data products:

- Experian **Clarity**
- Equifax **DataX**
- TransUnion **FactorTrust**

Covers **"alternative financial services"**, not in traditional credit report

- Online payday and installment loans, auto title loans, rent-to-own

What is the potential value of alt. data?

Lender's prior:

Consumers **with no credit information** (or thin file, short history) are **high credit risk**

Alt. data enables **lenders** to observe "shadow debt" (e.g., Argyle, Iverson, Nadauld, & Palmer, 2021) to update **priors on credit risk**

Potentially enabling more efficient lending

(expand lending, better price risk)

➡ **Applicants who would otherwise be rejected are riskier**, so higher auto loan interest rates
➡ **Probably efficient lending**, because hard to live in much of the U.S. without an auto

This Paper:
When lenders start using alt. data, how does this affect credit supply?

Data

- Loan origination data for all auto loans by 6 U.S. subprime auto lenders, 2014 to 2018.
- Traditional credit score (VantageScore) at loan application (& 2022)
- **"Hit"** indicator if in alternative credit data (**"alt. data"**)
- Alternative credit scores derived from alt. data
- Dates lenders adopt alt. data (+ loan applications post-lender adoption)

Methodologies

- **OLS & ML** for predicting auto loan delinquency and interest rates
- **Staggered difference-in-differences** with a continuous treatment (alternative credit score) exploiting variation in timing of alt. data adoption ➡ auto loan outcomes

When lenders start using alt. data, how does this affect credit supply?

Key Findings

- Alt. data predicts delinquency, beyond traditional credit scores

Subprime auto lender adoption leads to…

- Consumers ***in*** alt. data (“**Hits**”) interest rates ⬆
- Consumers ***not in*** alt. data (“**No-Hits**”) interest rates ⬇

When lenders start using alt. data, how does this affect credit supply?

Key Findings

- Alt. data predicts delinquency, beyond traditional credit scores

Subprime auto lender adoption leads to…

- Consumers ***in*** alt. data (“**Hits**”) interest rates ↑
- Consumers ***not in*** alt. data (“**No-Hits**”) interest rates ↓

Type of consumer that **selects into** using alternative credit products is ***more*** informative than other signals in alternative data (e.g., repayments)

When lenders start using alt. data, how does this affect credit supply?

Key Findings

- Alt. data predicts delinquency, beyond traditional credit scores

Subprime auto lender adoption leads to…

- Consumers ***in*** alt. data (**"Hits"**) interest rates ↑
- Consumers ***not in*** alt. data (**"No-Hits"**) interest rates ↓

Type of consumer that **selects into** using alternative credit products is ***more*** informative than other signals in alternative data (e.g., repayments)

(Presumably) lender profits ↑

How many consumers better-off vs. worse-off?

Paper argues ≈ all "Hits" worse-off

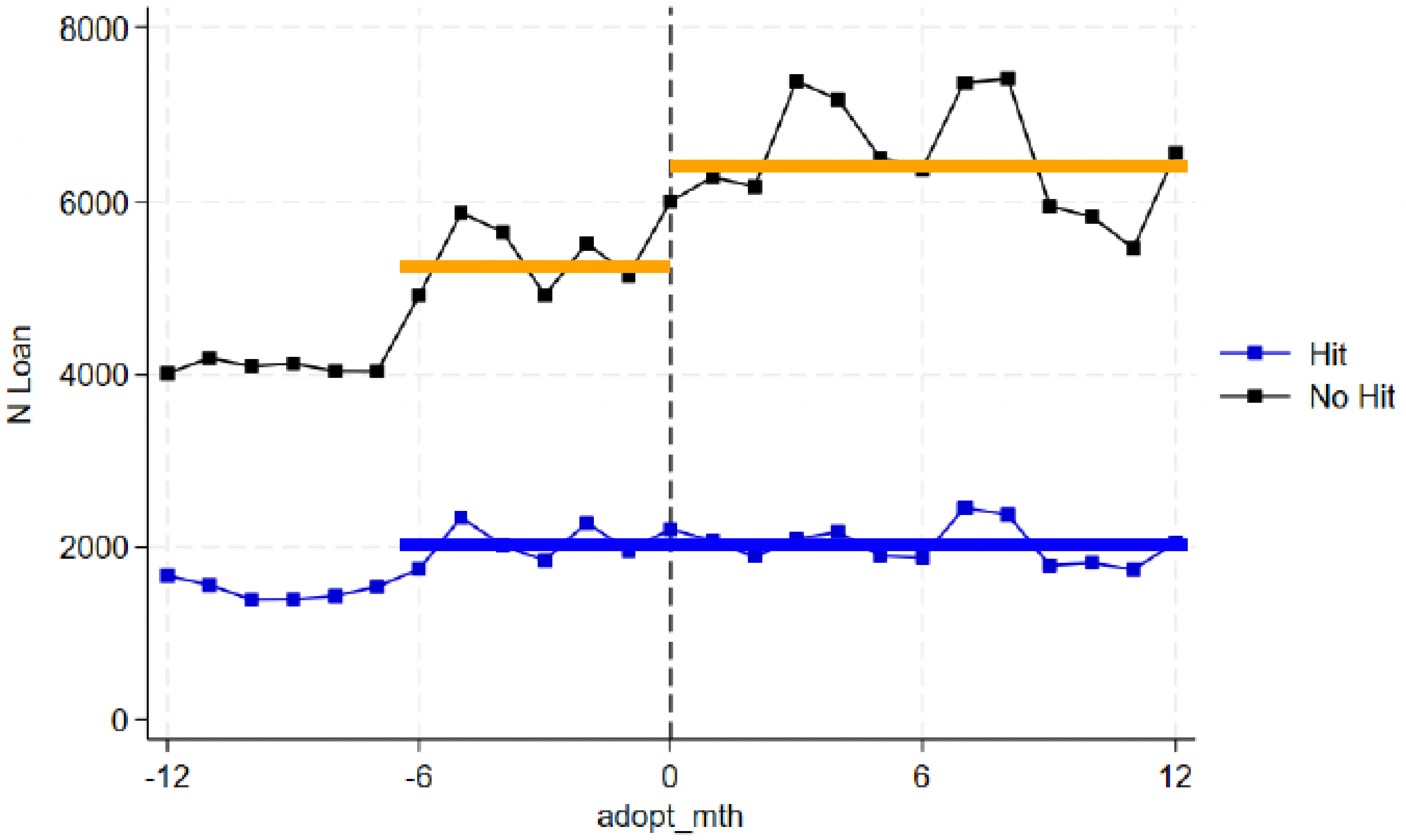
Extensive Margin

“Use of alternative credit data **limits credit availability** and **raises traditional loan costs** for most users of alternative financial services”

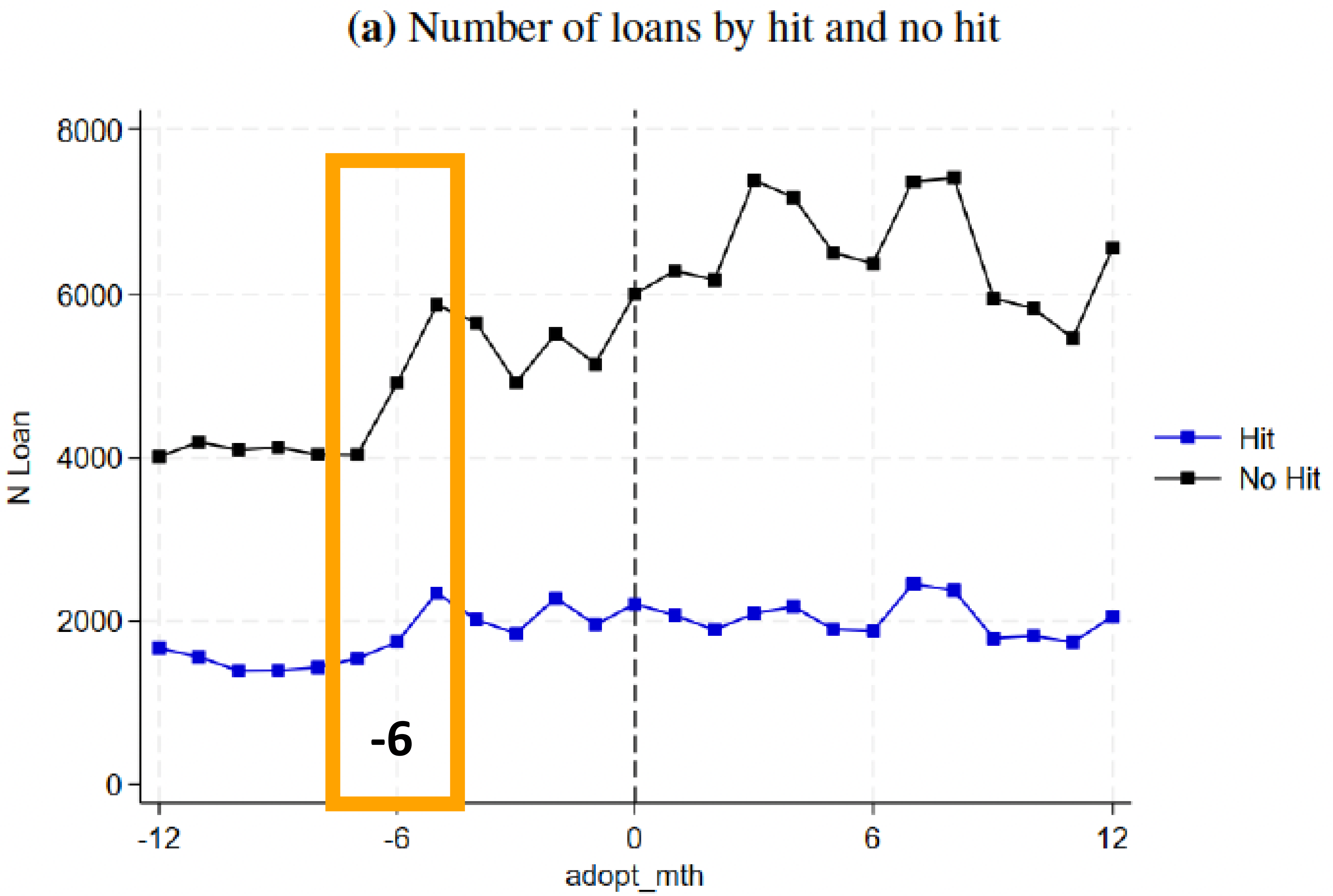
- I’m not convinced (yet!)
- Many potential margins of adjustment:
  - any loan, loan size, loan duration, loan interest rate, (car quality)

Large lending volume ⬆
~5,000 to 6,000+ loans per month to "No Hit" Consumers
No decline in lending to "Hit" Consumers

(a) Number of loans by hit and no hit

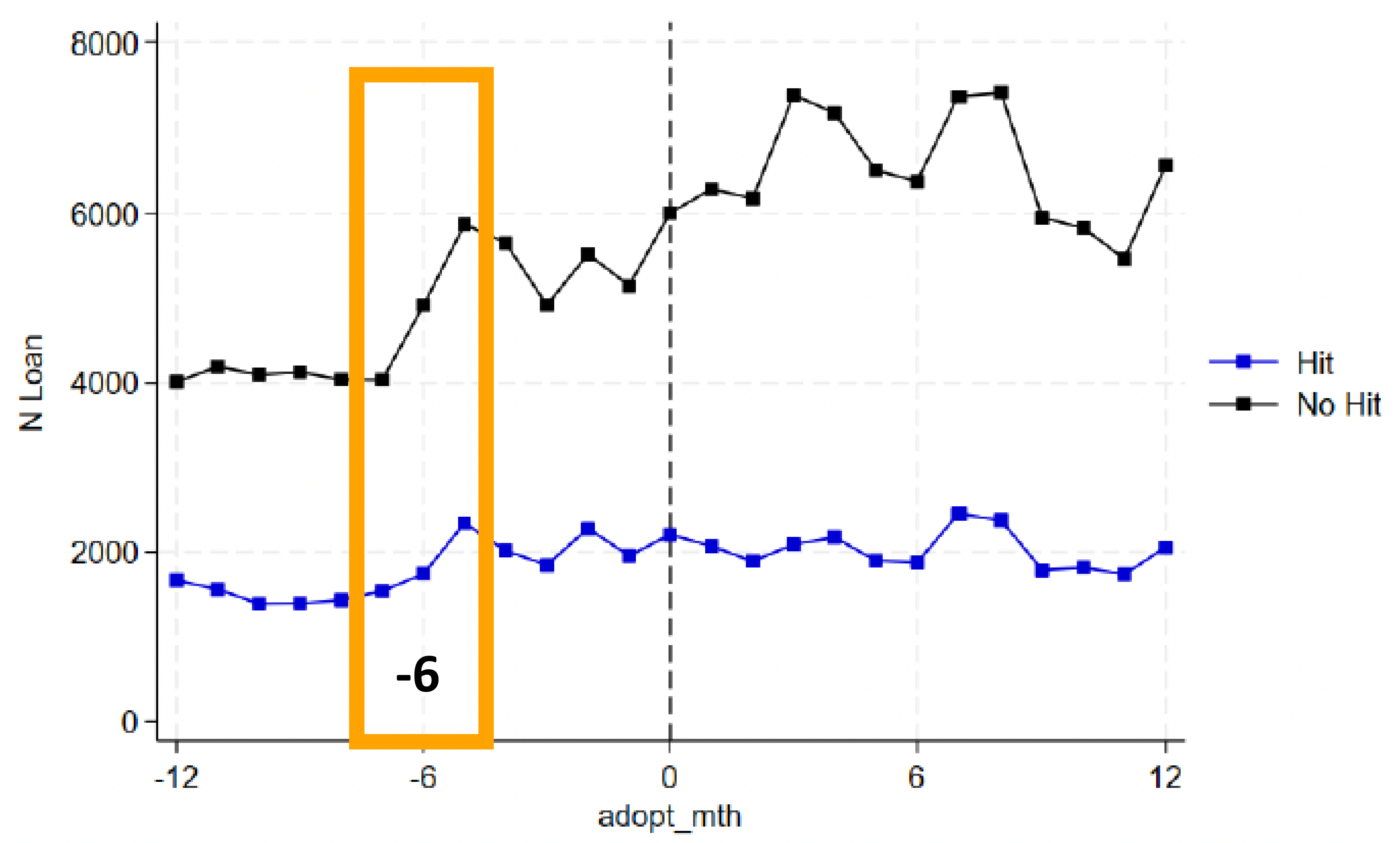


Why large ⬆ in “Hit” & “No Hit” loan volumes at t-6?



Why large ⬆ in "Hit” & “No Hit” loan volumes at t-6? Why stop at t+12? Interest rate results just emerging!

(a) Number of loans by hit and no hit



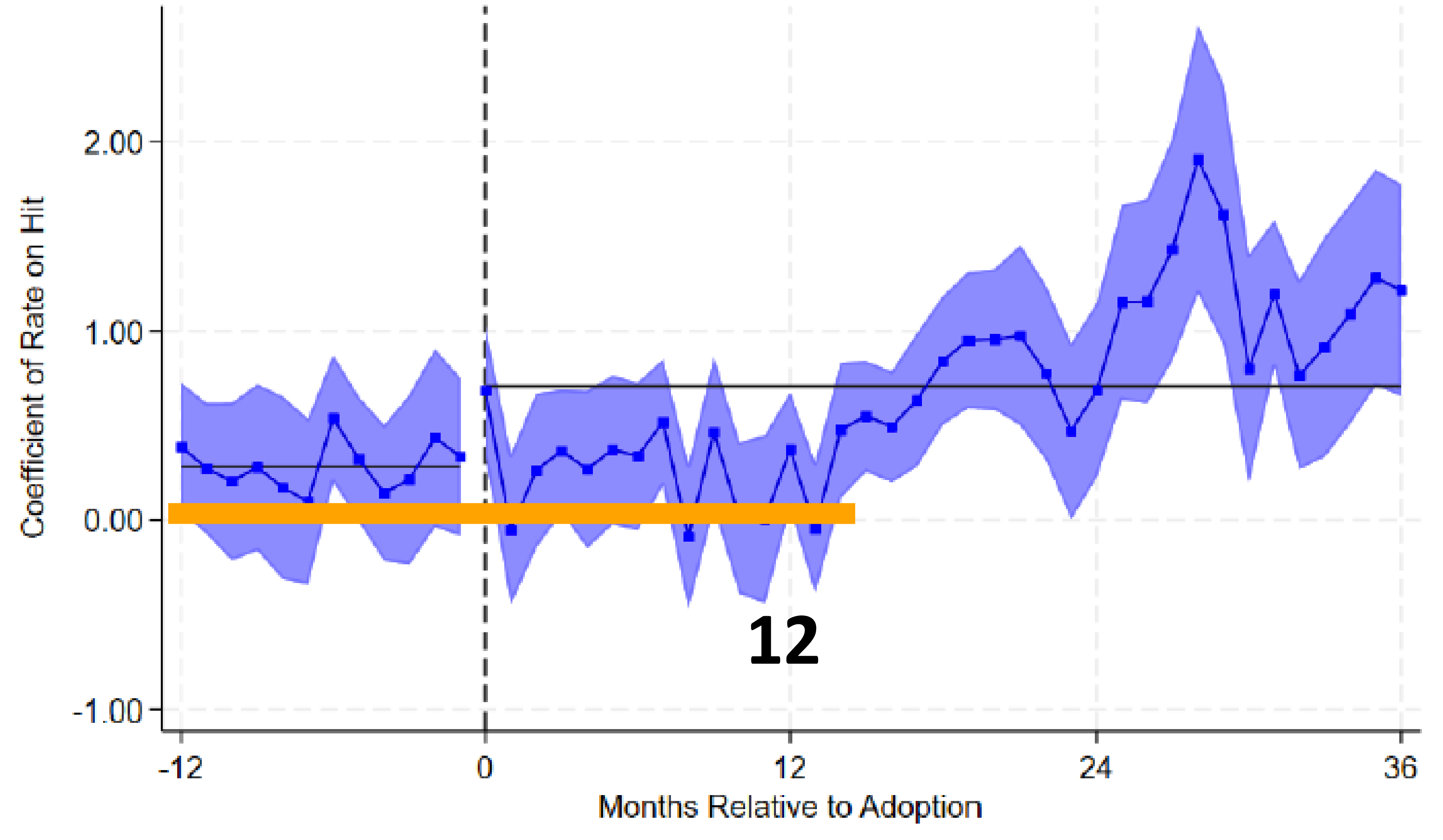


Figure 6. The Effect of Hit on Loan Rates Over Time. This figure plots the coefficients of a

Evaluating the key trade-offs of alt. data would need better data

1. What happened to **auto loan acceptance rates**?
    - Maybe you can get this information from one lender?

Evaluating the key trade-offs of alt. data would need better data

1. What happened to **auto loan acceptance rates**?
    - Maybe you can get this information from one lender?

2. Do **declined** applicants **get any auto loan elsewhere**? If so, what terms?
    - Solvable with credit reporting data

      (although buy here pay here auto loans not reported, e.g., Low, Clarkberg, & Gardner, 2021)

Evaluating the key trade-offs of alt. data would need better data

1. What happened to **auto loan acceptance rates**?
    - Maybe you can get this information from one lender?

2. Do **declined** applicants **get any auto loan elsewhere**? If so, what terms?
    - Solvable with credit reporting data
    (although buy here pay here auto loans not reported, e.g., Low, Clarkberg, & Gardner, 2021)

3. Does alt. data sharing **produce longer-term broader benefits**?
    - E.g., better loan terms / take out mortgages / faster credit score progression
    - Solvable with credit reporting data

Credit Information Market Design

What information is valuable for credit scoring?

Credit Scoring Model:

- Pr $(Y_{t+1} = 1) = f(X'_t\ \beta)$
- Look for features $(X'_t = [x_{1,t}\ ,\ x_{2,t}\ ,\ \ldots])$ that predict default (Y_{t+1})
- A valuable feature is an informative signal to **discriminate consumer types**:
  - Goods (non-defaulters) from bads (defaulters)
- Some features are deemed **unfair** & not allowed to be use (e.g., gender, race)

Which features predict default, & ⬇ credit score?

- Bankruptcy
- Defaulting in the Past
- High Credit Card Utilization

    -> Potentially ⬇ moral hazard

Which features predict default, & ⬇ credit score?

- Bankruptcy
- Defaulting in the Past
- High Credit Card Utilization

  -> Potentially ⬇ moral hazard

- Shopping around (applying for credit, closing old account)
- Moving address (UK credit scores not US FICO/VantageScore)

  -> Potentially deterring economically beneficial behaviors (searching/switching to better products, moving to opportunity)

Negative signals are valuable to lenders
Are better credit information market designs possible?

#1: Credit Building Loans

- Do not increase credit scores (Burke, Jamison, Karlan, Mihaly, & Zinman, 2023)
  - But they could do if scores recognized credit builder loans as a ‘different’ loan type

Negative signals are valuable to lenders
Are better credit information market designs possible?

#1: Credit Building Loans

- Do not increase credit scores (Burke, Jamison, Karlan, Mihaly, & Zinman, 2023)
  - But they could do if scores recognized credit builder loans as a ‘different’ loan type

#2: Buy Now Pay Later (BNPL):

- If you put BNPL into FICO, scores would ⬇
- FICO & VantageScore are adjusting models to account for BNPL being ‘different’
- BNPL lenders starting to report some loans to some bureaus (not in scores)

Negative signals are valuable to lenders
Are better credit information market designs possible?

#1: Credit Building Loans

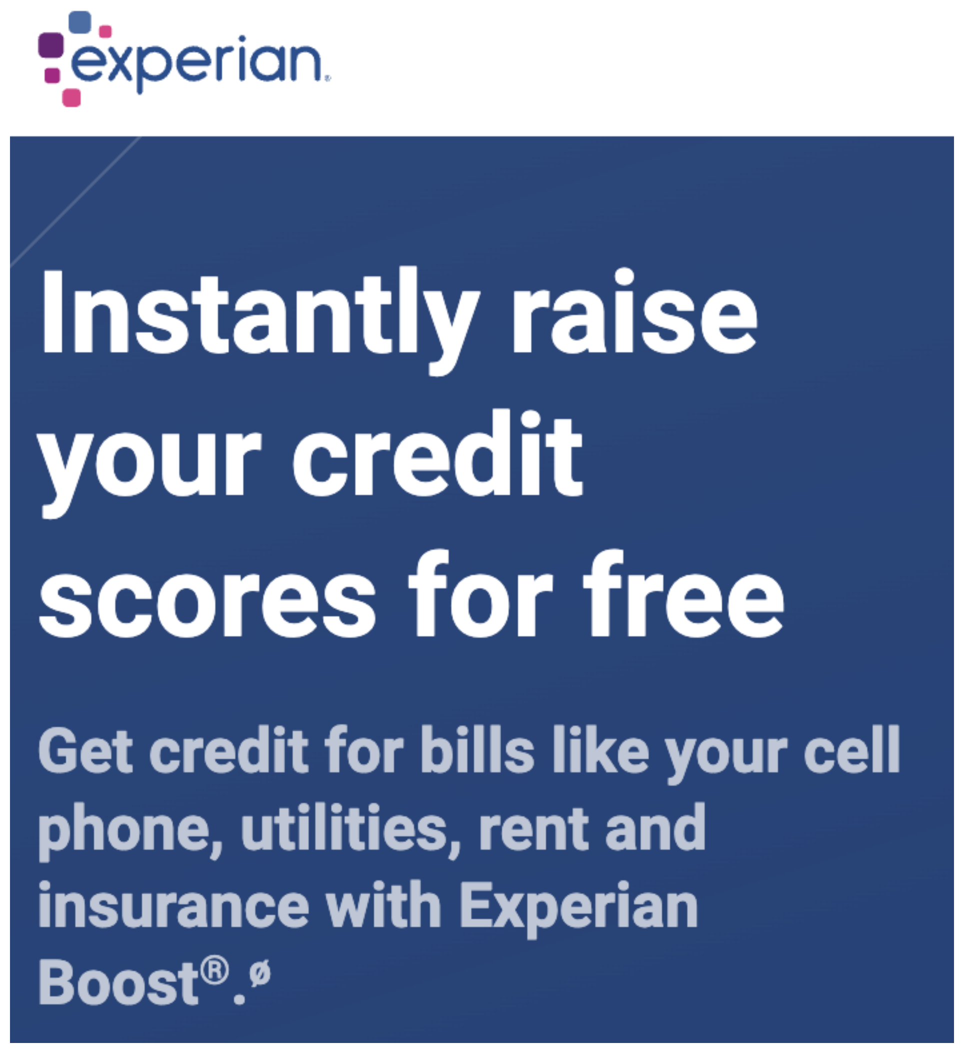
- Do not increase credit scores (Burke, Jamison, Karlan, Mihaly, & Zinman, 2023)
  - But they could do if scores recognized credit builder loans as a ‘different’ loan type

#2: Buy Now Pay Later (BNPL):

- If you put BNPL into FICO, scores would ⬇
- FICO & VantageScore are adjusting models to account for BNPL being ‘different’
- BNPL lenders starting to report some loans to some bureaus (not in scores)

#3: Experian Boost

- Consumers opt-in to select data to add
- By design, adding data cannot reduce Boost credit score



Thank You! “Borrowers in the Shadows: The Promise and Pitfalls of Alternative Credit Data”

“No sleep, no sleep until I’m done with finding the answer” (The Rasmus, 2003, In The Shadows)

www.benedictgk.com

benedictgk@rice.edu

@gk_ben / @profgk.bsky.social

The **marginal** value of information not in traditional credit reports

I really want to know the **marginal impacts** of hit and alt. credit score relative to model that doesn't use these sources of information.

- SHAP values go some way towards this
- But how does adding this information affect AUC (or other performance metrics)? A simple table comparing performance metrics across 4 models (baseline model using VantageScore and other variables used in underwriting pre-adoption, baseline+hit, baseline+alt.score, baseline+hit+alt.score) would be very informative. Follow standard approach used in other papers, Then show how the distribution of the changes in the probabilities of default (segment by hit) to show 'winners' and 'losers'?

(e.g., Berg, Burg, Gombović, & Puri, 2020; Fuster, Goldsmith-Pinkham, Ramadorai, & Walter, 2022; Blattner & Nelson, 2025; Duarte, Fonseca, Kohli, & Reif, 2025; Guttman-Kenney, 2025)

You should discuss Blattner & Nelson (2025). They show very limited value of alternative data. The reason is that alt. data borrowers are usually already distressed in main credit file, so there isn't much new information. One potential difference is your sample of consumers.

Many papers show the value of non-traditional data sources for predicting defaults. I think it would be good to link to some of these (you don't need to cite them all!) and compare your estimates. You should link to this literature. The key strength of your data over some of these studies is that you can observe the supply response of lenders' contract terms.

- Checking account data, increasingly via open banking (e.g.,Berg, Burg, Gombović, & Puri, 2020; Alok, Ghosh, Kulkarni, & Puri, 2025; Chioda, Gertler, Higgins, & Medina, 2025; Babina, Bahaj, Buchak, De Marco, Foulis, Gornall, Mazzola, & Yu, 2025; He, Huang, & Zhou, 2023)
- Rental data (Thompson Cochran & Stegman, 2022; Theodos, Teles, Lieberman, 2025)
- Mobile phone behaviors (Björkengren & Grissen, 2020)
- Grocery data (Lee, Yang, & Anderson, 2025)
- Buy now pay later (Laudenbach, Molin, Roszbach, & Sondershaus, 2025)
- More examples in appendix of our JEL (Gibbs, Guttman-Kenney, Lee, Nelson, & van der Klaauw, & Wang 2025)

Econometrics issues

- See the ‘new’ difference-in-differences literature for how to estimate effects when you have staggered treatment. The current draft does not even mention these known issues. I worry that your current estimates, with effects appearing 12+ months after adoption, may be driven by how the data and estimates have been set up.
- In addition, you are using continuous treatment effect. This relies on even stronger assumptions (see Callaway, Goodman-Bacon, & Sant’Anna, 2024). An easy solution would be to not using a continuous treatment effect?
- Showing a descriptive time series of the key outcomes for the lenders would be informative. I want to see this for all the potential margins of adjustment. We should see jumps at adoption dates. This would enable the reader to evaluate how good the control groups are, and then you can convince us further with regressions. Table A.1 shows quite large differences in levels across the 6 lenders. Clarkberg, Gardner, & Low (2021) shows that there are large differences across lenders so I worry about how suitable your control groups these are. Different lenders may have different incentives to share data, so the ordering of adoption may matter.
- Standard errors should be clustered at the source of variation: lender-level. Unfortunately doing so will substantially affect your power. You only have 5-6 lenders so I suggest you look at small-sample clustering methods. See Alvarez, Ferman, Wuthrich (2025) for the cutting-edge (and Conley & Taber, 2011, RESTAT).
- I didn’t follow why you include county F.E. separately. If you think county is important, it would make sense to include lender-by-month**-by-county** F.E.
- You don’t say what your control variables are in your diff-in-diff regression.
- Are consumers without VantageScores getting dropped in some of your regressions? I think that you need an indicator for no VantageScore to allow them to be different.
- (minor) People don’t normally use OLS to predict delinquency because it is highly non-linear. Logit or ML methods would align closer to what industry actually does and the literature (cited on previous page).
- (minor) In Table 3,4,5 why sis N so much lower when predicting rates compared to when predicting delinquency. Maybe keep consistent sample?

Current draft needs to make clearer the institutional details on how lenders use alternative score & hits.

I need more details to evaluate which loan applicants this alt. data information is affecting, and how. Some questions I had.

- Are there sequential steps? If application would get declined in normal underwriting, only then go to Alternative Score. Are they using ‘hits’ as well as the score? At the same time?
- Is it used for both decision to lend and pricing decision?
- (p.12) “*After adoption, the coefficients increase steadily, suggesting that lenders progressively incorporate the information from alternative credit data, leading to higher interest rates for borrowers included in the data.”* <- I didn’t see any steady increases. There’s nothing until at least 12 months post-adoption. This suggests that lenders may have been purchasing the data but not using it for lending decisions. It casts doubt over your adoption dates, and makes me concerned that your effects may be driven by other changes at the lenders (or that the effects are driven by econometrics issues or the control group not being comparable).
- With only 6 lenders you could you work out their discontinuous lending decisions / interest rate setting rules and exactly how changed after using alternative data. What thresholds change? What information is getting more/less weight in the decision? Pinning down exactly how lenders adjust their lending criteria is hard to observe, so is a strength of your data.

Miscellaneous Comments

- Your abstract and introduction should make clearer that your outcomes are auto loan delinquency and auto loan interest rates. When reading it wasn't until the data section that it became clear that you aren't looking at overall delinquency rates.
- Section 3, your XGBoost model's AUC of 0.679 appears pretty poor compared to prior studies. Blattner & Nelson (2025) and Chioda, Gertler, Higgins & Medina (2025) are probably the closest comparison to your setting. It would be good to understand why this is compared to other papers. Also normally presenting other performance metrics as robustness is helpful., as is having a benchmark saying what a naïve model that predicts everyone defaults would do.
- Ghosts. What % are these would be informative to include in Table 1? In Figure 9, what happens to SHAP value for these? What is the overlap between 'hits' and 'ghosts'? Alt. data may help ghosts the most, by demonstrating they they are not in the alt. data so less high risk. Showing the time series of loan volumes for ghosts seems very important! You could also split this by hit & no hit.
- It wasn't clear from your data section whether your data are linked across lenders. It would be very informative if you can follow a consumer over time. It also wasn't clear whether your July 2022 data includes other information, such as auto loans on their credit report. A single snapshot of credit reporting data at the tradeline—level includes accounts open and closed so would be sufficient to give you a history of auto loan originations, their origination dates, loan terms, and you can estimate interest rates (see my JEL).
- You setting seems well-suited to using Jansen, Nagel, Yannelis, & Zhang framework to speak to welfare? If not, explain why not.
- You use standardized credit scores in your regressions but non-standardized credit scores in your figures. This makes the results hard to compare.
- Figure 10 notes refer to a red line but there was none I could see.
- Figure 11 panel b was difficult understand without clearer labels
- You can compare your Figure 3 to Figure A.5 in Blattner & Nelson (2025) showing FactorTrust Match Rates by VantageScore
- Footnote 1. There are more direct & recent examples on this point. see Blattner & Nelson (2025) and Di Maggio & Ratnadiwakara (2024).
- Footnote 8. Wasn't clear why you take a different approach to the lender here. Suggest following the lender's approach, or explaining why not.
- It would be informative to show how the slopes in Figures 3 and 4 vary post adoption.
- When predicting delinquency and interest rates, don't use information such as race that a lender cannot use in underwriting.
- My understanding of alt. data (caveat: I've not worked with them) is that data coverage / quality substantially varies over time, and there is limited performance data which may explain why there is limited information beyond the hit rate.