### Importance of access to credit:

Access to credit has long been recognized as a critical driver of economic growth, enabling individuals to seize opportunities for upward mobility.

Conversely, the exclusion from financial access can have profound implications, particularly for marginalized communities, leading to a perpetuation of poverty and limited avenues for progress. The detrimental effects of unequal credit access based on factors such as race and gender are especially noteworthy, as they exacerbate the existing disparities, leaving previously disadvantaged groups further crippled and deepening the gaps in welfare and opportunity (King and Levine, 1­993).

The link between credit accessibility and economic growth has been a subject of interest for researchers and policymakers alike. A functioning and inclusive financial system plays a vital role in providing the necessary capital to finance entrepreneurial projects, investments in human capital, and infrastructure development, which are all instrumental in fueling economic advancement. Studies have revealed that access to credit can substantially enhance productivity, facilitate innovation, and encourage risk-taking, leading to a multiplier effect on economic growth and overall prosperity (Beck et al., 2007).

The credit allocation decision-making process involves evaluating the creditworthiness and risk profile of potential borrowers to determine whether they qualify for the requested credit and, if so, under what terms and conditions. For this purpose, several aspects are considered, and the process must comply with legal and regulatory requirements to ensure that lenders adhere to fair lending practices to prevent discrimination based on factors such as race, gender, nationality, or religion.

### EXPLANATION OF ALGORITHMIC APPROACH TO CREDIT-THINGS PREDICTION + HISTORY

Over the past three decades, consumer lending has witnessed a significant shift from traditional interview-based underwriting to data-driven models for evaluating and pricing credit risk. In the early 1990s, most lenders relied on a single "house rate" and conducted borrower interviews for screening purposes (Johnson, 1992).

Soon after, processes started to become automated, and statistical methods have been used consistently in various applications. The first notable example of the statistics applied to credit risk analysis can be traced back to 1959, when Fisher built a model using ordinary least squares (OLS) to explain the determinants of a bond’s risk premium.

The introduction of credit scoring, together with the computerization of the industry, has increased the availability and affordability of credit, by allowing lenders to assess more quickly the creditworthiness of applicants. At the same time, credit scoring is also a powerful instrument that can make evaluation more objective and prevent discrimination, which has negative effects both on fairness and on economic outcomes by misallocating available resources (Avery, 2009).

Generally, it is fair to say that efforts to predict credit-related variables using statistical methods have been quite successful over the years, and with the decline in storage and computing costs, lenders progressively adopted estimates of default risk to determine individual loan prices. ADD EXAMPLES. A major advantage of employing such techniques is that statistical methods are succinct and easy to explain, and yet bring quite satisfactory results. However, they are theoretically not suited to work with financial data, which is generally not compliant with standard assumptions, including multivariate normality assumptions (Huang et al., 2003).

More recently, technological advancements have allowed for the implementation of artificial intelligence methods capable of automatically extracting knowledge from data sets. This enables the learning of the model's specific structure directly from the data, thus increasing the explanatory power of the model itself. However, this leads to the emergence of complex and hard-to-explain models, shining a new light to the trade-off between a model's explanatory power and parsimony, in which the first enhances prediction accuracy, while the latter ensures the model's generalizability and interpretability (Huang et al., 2003).

### FINTECH + new digital footprint –(connettere con AI methods)

This increased predictive power of new technologies has more recently been combined with new available sets of data, giving rise to FinTech, described by Schueffel (2017) as “a new financial industry that applies technology to improve financial activities”.

FinTech has given a push the whole financial industry, prompting a general innovation and modernization trend to maintain competitiveness.

For example, FinTech lenders can be distinguished from traditional lenders based on their efficiency. They have been shown to have lower default rates on comparable loans, while being 20% faster than their counterparts at processing mortgage applications. As mentioned, this is due to the use of more sophisticated algorithms for credit scoring and pricing of loans that are better able to fit and learn from the underlying data (Fuster, 2019).

Additionally, researchers found that fintech lenders seem to use different information for credit-related decisions than non-fintech lenders, and this can be likely traced back to using different dimensions of “big” data that not available to traditional lenders. More specifically, these variables include the so-called “digital footprint”, which refers to the trail of data and information that individuals create and leave behind when using digital devices and the internet, ranging from the type of device and operating system used, the email and the behavior on a website (Buchak et al., 2018) to location, payment, and social media data (Koren, 2016).

These variables, together with more traditional information, have been shown to be useful for predicting default without decreasing the total number of accepted credit requests. This is due to a positive reshuffling effect, by which customers with favorable digital footprints and low credit score gain access to credit, while customers with unfavorable digital footprints and medium credit scores lose access. This effectively implies that lenders who can complement traditional credit bureau information with digital footprints will be able to make superior lending decisions (Berg, 2020).

Additionally, research shows that the informational content of digital footprint data is better modeled for prediction of borrowers' losses and defaults by machine learning techniques than traditional statistical methods.

This difference becomes even more notable in the presence of an external shock to the aggregate credit supply, where the improved performance of machine learning models compared to traditional methods can be pinpointed to the better ability to mine and model non-linear relations within the data (Gambacorta et al., 2019) .

### G20 High-Level Principles for Digital Financial Inclusion

The unique potential of the digital footprint lies in the widespread availability of smartphones users around the world, including the billions of adults who still lack access to the traditional financial services, which makes it a powerful tool for promoting financial inclusion. Since it effectively serves both scorable and unscorable customers, the use of digital footprints thus has the potential to boost access to credit and lower geographical, racial and gender inequality.

As a matter of fact, the G20 High-Level Principles for Digital Financial Inclusion (2016), along with international standard-setting bodies' principles, provide a basis for country-specific action plans to leverage the potential of digital technologies. G20 members endorse these principles and encourage countries to incorporate them into their broader financial inclusion strategies, particularly in the realm of digital financial inclusion. The goal is to facilitate concrete and meaningful actions in advancing financial inclusion with the aid of digital financial services.

In recent times, the FinTech industry has shown a keen interest in harnessing this potential and is challenging conventional banking practices and devising inventive financing solutions, motivated also by the significant rise in accessible digital footprints in developing economies (Kendall, 2017). Provided that consumer interests are safeguarded, and concerns related to privacy, security, and ethical usage are adequately addressed, utilizing this data could prove to be a valuable approach in reaching unbanked individuals from economically disadvantaged backgrounds, offering them a diverse array of financial products (Kumar and Muhota, 2012).

### Discrimination

As a matter of fact, it is essential to consider the US legal framework and analyze the impact of discrimination on US credit markets. For this purpose, the Home Mortgage Disclosure Act (HMDA) requires lenders to collect and disclose individual-level data, including race and gender, aiming to foster transparency, accountability, and academic scrutiny in a domain with a history of contentious discriminatory practices. Economists have consistently found evidence of racial discrimination (CITARE QUALUNO), although some argue that disparities in outcomes may be attributed to differences in default risk (Ladd, 1998).

Additionally, Blattner and Nelson (2021) uncovered a significant information disparity faced by lenders in the US credit markets when evaluating default risk among historically under-served groups. Widely used credit scores are shown to exhibit higher statistical uncertainty for these groups, and the researchers find that equalizing the precision of credit scores can lead to a reduction of approximately half in approval rate disparities and credit misallocation for disadvantaged groups. These findings emphasize the crucial role of addressing information disparities to promote fair and efficient credit market outcomes (Blattner and Nelson, 2021).

In the United States, credit risk evaluation operates under strict regulations and has not fully embraced algorithmic decision-making. However, even in mortgage lending, researchers are exploring the potential of machine learning algorithms to improve default predictions and enhance financial inclusion for those who might have been excluded under simpler decision-making processes (Fuster et al., 2019).

These techniques shine a new light on unfair price discrimination and equity, raising persistent policy concerns about equity across consumers (Avery et al., 2009; Traub, 2013). As machine learning advances, it becomes essential to address these issues to ensure fair and equitable credit access for all consumers. In fact, algorithmic decisions are often less transparent and harder to explain than a straightforward rules-based process, and although both human decision-making and rules-based processes are susceptible to unfair biases and inaccuracies, algorithms face heightened scrutiny due to their reduced transparency and potential for broad scalability. While a human decision-maker may exhibit varying judgments influenced by cognitive biases, an algorithmic decision based on bias can perpetuate discrimination on a larger scale, raising concerns about fairness and accountability (Lee and Floridi, 2020).

Even when these algorithms are not explicitly fed with protected characteristics such as race, religion, gender, or disability, they can indirectly infer such information, leading to disparate treatment. Research on US mortgages has revealed that Black and Hispanic borrowers are disproportionately less likely to benefit from the implementation of machine learning in credit scoring models (Fuster et al., 2019). Moreover, recent studies have cautioned that flexible statistical technologies like machine learning may reduce overall loan approval rates for disadvantaged groups (Fuster et al., 2020) and continue to generate cross-group disparities in loan terms in FinTech underwriting (Bartlett et al., 2022). Additionally, credit scores have been linked to geographic misallocation in the US mortgage market (Hurst et al.2016), and important associations have been found between geography and interest rates paid by minority borrowers (Bartlett et al., 2022).

It is therefore of the utmost importance to isolate those variables which might become sources of discrimination, even though they might not appear to be related with any sensitive characteristic such as race or gender. This would ensure that minorities are not subject to discrimination in the form of disparate impact, a concept which centers on outcomes rather than intent, encompassing any policy or practice that disproportionately disadvantages a protected group, regardless of the intention to discriminate (Baum et al. 2015).

### 6. focus + RQ

This thesis aims to investigate the performance of machine learning algorithms and their influence on discrimination in credit allocation, with a particular focus on the mortgage market. This market significantly contributes to the continuation of wealth disparities between generations, especially for historically disadvantaged minorities, who are less inclined to invest into homeownership and accumulate home equity (Charles and Hurst, 2003; Kuhn et al., 2020; Charles and Hurst, 2002). The central question addressed is whether machine learning techniques can enhance fairness and mitigate biases in credit allocation.

Moreover, there exists a literature gap in analyzing and comparing regression techniques for predicting interest rates, as most studies have primarily focused on classification problems such as default or credit scores. This research aims to bridge this gap by exploring and contrasting regression techniques' effectiveness in interest rate prediction.

### VECCHIO FOCUS OF THESIS: attention to mortgage (CONNETTERE)

The focus of this thesis will be around the performance of machine learning algorithms and the impact they have on discrimination in credit allocation. Particular attention will be directed specifically towards the mortgage market, which significantly contributes to the continuation of wealth disparities between generations (Charles and Hurst, 2003; Kuhn et al., 2020), especially in the context of historically disadvantaged minorities, who are less inclined to invest into homeownership and accumulate home equity (Charles and Hurst, 2002).

The

### RESEARCH QUESTIONS connettere sopra

Specificare che c’è un gap nella letteratura nell’analizzare e confrontare tecniche di REGRESSIONE per predirre interest rate e non tecniche di classificazione che sono state ampiamente comparate per predirre default sui debiti o credit scores.

* + the detailed research question/s addressed in the thesis and the logical reasoning behind their development);

ALTRO :

* **Frost et al (2019)**: when it comes to predicting loss rates, credit scoring techniques based on big data and machine learning have so far outperformed credit bureau ratings. A key question here is whether this outperformance will persist through a full business and financial cycle. new forms of non-prudent risk-taking???
* Possible literature for conclusion for why minorities do not actively look for better rates (from Bartlett 2022):

Woodward, S.E., 2008. A Study of Closing Costs for FHA Mortgages. Re-

port. U.S. Department of Housing and Urban Development. Woodward, S.E., Hall, R.E., 2012. Diagnosing consumer confusion and sub- -optimal shopping effort: theory and mortgage-market evidence. Am.

Econ. Rev. 102, 3249–3276.

RIGUARDARE PAPER BARTLETT ? Bartlett 2022

Gambacorta – NON DISCRIMINATION!!

1. The **predictive power of all the models improves when the length of the relationship** between bank and customer increases. However, the comparative advantage of the model that uses the fintech credit scoring technique based on machine learning tends to decline when the length of the relationship increases.

**6. Credit scoring and relationship lending**

Traditional financial system: relationship-based, trust and human interaction

Fintech: transaction-based, no long-term relationship with the customer, typically short-term credit lines that can be automatically cut if a customer’s condition deteriorates

* performance of the three models improves with the length of the relationship
* Fintech comparative advantage seems to increase for low levels of the bank-customer relationship, but stays constant after a while, once the relationship has solidified

# PAPER LETTI

## Bartlett 2022 - Consumer-lending discrimination in the FinTech Era – DISCRIMINATION FOUND !

* "literature":

1. Buchak et al. (2018) find that “Relative to non- fintech shadow banks, fintech lenders... appear to use different information in setting interest rates, consistent with a big data component of technology.”
2. Fuster et al. (2019) study FinTech lenders in detail and conclude that the main difference between them and other lenders is efficiency: FinTech lenders process mortgage applications 20% faster. However, they also find that “Fin- Tech default rates are about 25% lower than those for traditional lenders, even when controlling for detailed loan characteristics.” Although they interpret this finding as evidence that FinTech lenders are not more lax in their screening than traditional lenders, it may also indicate the use of more sophisticated credit-screening or pricing algorithms.

* Results: **DISCRIMINATION FOUND !**:
* All the coefficients are still significantly greater than zero (minority borrowers have higher rates)
* Just FHA purchase loans have a significantly lower coefficient of FinTech x Minority than Non-FinTech x Minority, but there is not stat sign difference otherwise (see Table 4) -> FinTech does not reduce discrimination overall
* Vedere altri risultati nei dettagli (5.1 – 6.)

1. the results for GSE and FHA purchase lending in the 2018–2019 vintage loans is consistent with **FinTech lenders using pricing strategies and data analytics that produce discriminatory pricing.**

even if algorithmic lending can reduce discrimination relative to face-to-face lenders, it is insufficient to eliminate discrimination in loan pricing.

1. important **association between minority rate disparities and geography**

🡪  in the GSE and FHA markets : PAYING MORE (+ MORE LIKELY TO BE REJECTED)

## Gambacorta: How do machine learning and non-traditional data affect credit scoring?

The main conclusions of our paper can be summarised as follows:

1. The fintech’s **machine learning-based credit scoring models outperform traditional** empirical models (using both traditional and non-traditional information) in **predicting borrowers’ losses and defaults.** 
   1. Credit card data (what the bank usually observes) + access to apps and platform usage data
2. **Non-traditional information improves** the predictive power of the model.
   1. In terms of contribution of non-traditional data and machine learning to predictive power, non-traditional data contribute an additional 2.2% of the AUROC (=(0.607-0.5939)/0.5939), while applying machine learning techniques provides an additional 5.3% of the AUROC (=(0.6391-0.607)/0.607).
3. While the models perform similarly well in normal times, the model based on machine learning is better able to predict losses and defaults following a negative shock to the aggregate credit supply. One possible reason for this is that machine learning can **better mine the non-linear relationship** between variables in the event of a shock (regulatory shock).
4. The **predictive power of all the models improves when the length of the relationship** between bank and customer increases. However, the comparative advantage of the model that uses the fintech credit scoring technique based on machine learning tends to decline when the length of the relationship increases.

**6. Credit scoring and relationship lending**

Traditional financial system: relationship-based, trust and human interaction

Fintech: transaction-based, no long-term relationship with the customer, typically short-term credit lines that can be automatically cut if a customer’s condition deteriorates

* performance of the three models improves with the length of the relationship
* Fintech comparative advantage seems to increase for low levels of the bank-customer relationship, but stays constant after a while, once the relationship has solidified

## **Blattner and Nelson** – già riformulato - DISCRIMINATION

Blattner and Nelson (2021) uncover a significant information disparity faced by lenders when evaluating default risk among historically under-served groups in the US credit markets. Widely used credit scores exhibit higher statistical uncertainty for these groups, primarily due to characteristics of the underlying credit report data rather than model or algorithm issues. By employing a structural model of lending that considers information heterogeneity, they quantify the impact of addressing these disparities in the US mortgage market and they find that equalizing the precision of credit scores can lead to a reduction of approximately half in approval rate disparities and credit misallocation for disadvantaged groups. These findings emphasize the crucial role of addressing information disparities to promote fair and efficient credit market outcomes (Blattner and Nelson, 2021).

## Avery 2009 **-** Credit Scoring and Its Effects on the Availability and Affordability of Credit

* **Credit scoring in general**: credit scoring has increased the availability and affordability of credit
  + Quicker assessment of creditworthiness for lenders
  + Possibly made evaluation more objective and less based on personal characteristics prohibited by law (eg race or ethnicity)

1. *Credit Scores Differ across Groups:* blacks, Hispanics, single individuals, those younger than age 30 and the individuals residing in low-income or predominately minority census tracts had lower credit scores on average
2. *Score Differences Reflect Differences in the Content of Credit Records:* Groups with lower average scores tended to experience a higher incidence of payment problems on credit obligations (eg higher % of blacks experiencing serious delinquency)
3. *Credit Scores Predict the Risk of Default*
4. *Access to Credit Differs Depending on Credit Score:* interest rates and denial rates decrease as credit score goes up

## WANG 2022 – **MODELLI** & METODOLOGIA

CONCLUSION

In this paper, we considered the mortgage rate prediction problem. The classical multiple linear regression is usually suboptimal due to violation of normality and presence of outliers. Three robust regression methods and deep neural networks are suggested and compared on two real data sets. Deep neural network is shown to make better prediction. It not only gives the minimal mean absolute error, but also has a mean residual closest to zero in most situations. Therefore, deep neural network is recommended for practice use by mortgage companies or banks.

Although default prediction provides critical information for the mortgage companies to make decisions, they do not directly help determine the appropriate interest rates

**RIFORMULARE PER ME PER SPIEGARE IL GAP NELLA LITERATURE !!!**

While it takes a variety of risk factors into account and adjusts for the selection bias, the prediction of mortgage rates for approved applicants is still based on linear regression models. As the great success of modern machine learning and artificial intelligence in data processing and predictive analytics is triggering significant changes in human life, business models, and industries, especially in the fields of automation, robotics, and online sales, to our best knowledge, their adoption in mortgage rate prediction seems sparse. In recent literature some machine learning approaches such as boosted regression tree, random forests, and convolutional neural networks were used for mortgage default prediction (Fitzpatrick & Mues, 2016; Kvamme et al., 2018). Although default prediction provides critical information for the mortgage companies to make decisions, they do not directly help determine the appropriate interest rates.

Data: ?????

## Lee, Floridi

* Both human decision-making and rules-based processes are susceptible to unfair biases and inaccuracies. However, algorithms face heightened scrutiny due to their reduced transparency and potential for broad scalability. Unlike a straightforward rules-based process, algorithmic decisions are often less transparent and harder to explain. While a human decision-maker may exhibit varying judgments influenced by cognitive biases, an algorithmic decision based on bias can perpetuate discrimination on a larger scale, raising concerns about fairness and accountability. 🡪 già riformulato
* propose a new methodology that focuses on relational trade-offs rather than absolute mathematical conditions
* Far riscrivere a chatgpt 2.1 framework SU **DISCRIMINATION**
  + Spiega differenza tra intenzionale e non ma che ha cmq impatto
* **PRENDERE DA ONE NOTE !!!**
  + **Modelli**
  + **Trade off**

MODELLI:

TODO: Questione FAIRNESS:

* Escludere sensitive vars dai modelli e poi rerunnarli con anche sensitive
* Fare log reg di tutte le variabili nel dataset per predire Minority, per spiegare questione "5.3 PROXIES OF RACE" (Lee, Fuster) 🡪 copiare modalità di spiegazione

## SINGH CAPIRE PER METRICA

## HUANG 2004 - seminal work - FINITO

bond-rating ALGORITHMIC prediction problem

* **STAT METHODS:**

The use of statistical methods for bond-rating prediction can be traced back to 1959, when Fisher utilized ordinary least squares (OLS) in an attempt to explain the variance of a bond’s risk premium […]

The general conclusion from these efforts in bond-rating prediction using statistical methods was that a simple model with a small list of financial variables could classify about two-thirds of a holdout sample of bonds. Methods are succinct and were easy to explain, but theoretically not suited with financial data, which is generally not compliant with multivariate normality assumptions

* **AI METHODS:**

automatically extract knowledge from a data set and construct different model representations to explain the data set, allow learning the particular structure of the model from the data --> usually very complicated and hard to explain; trade-off between the explanatory power and parsimony of a model, where explanatory power leads to high prediction accuracy and parsimony usually assures generalizability and interpretability of the model

[… esempi di paper che usano ML ma solo per classification …]

SCRIVERE QUESTA COSA MA AL CONTRARIO:

Some other researchers have studied the problems of default prediction and bankruptcy prediction [26,41,49], which are closely related to the bond- rating prediction problem. Similar financial variables and methods were used in such studies and the prediction performance was typically higher because of the binary output categories. Forse già messo? bo

## Berg 2020: **On the Rise of FinTechs: Credit Scoring Using Digital Footprints** - FINITO

1. **discussing the key economic outcomes and implications of our findings** 
   1. decompose the explanatory power of the digital footprint into each of the individual variables --> all significant together
   2. default rates drop significantly after the introduction of the digital footprint, but the proportion of customers with access to credit stays roughly the same -> good footprint & low credit score gain access, medium score & bad footprint lose access
   3. Check that digital footprint works well for both scorable and unscorable customers -> can boost financial inclusion
   4. Discuss implications of findings for customers, firms and regulators' behaviors.

* Results: default rate in a multivariate setup using a simple time series difference (default rates post- vs. pre-introduction of the digital footprint) : Default rates significantly drop by approximately 50% around October 19, 2015, whereas the number of purchases made via invoice remains unchanged. This figure suggests that using more information (adding the digital footprint for all observations and adding the credit bureau score for some of the observations) helped to significantly reduce default rates. It also highlights a reshuffling effect, as opposed to a simple explanation or contraction effect: customers with favorable digital footprints gain credit access while customers with unfavorable digital footprints lose credit access.
* Digital footprint unique bc available for virtually everyone
* Still, recent activity in the FinTech industry suggests this is an avenue that FinTechs aim to take. Motivated by a dramatic increase in the availability of digital footprints in developing economies, **new FinTech players have emerged that use digital footprints to challenge traditional banking options and develop innovative financing solutions**. (See, for example Kendall (2017) )

1. **Conclusion**

* show that even simple, easily accessible variables from the digital footprint match the information content of credit bureau scores
* digital footprints complement rather than substitute for credit bureau information, implying that a lender that uses information from both sources (credit bureau + digital footprint) can make superior lending decisions
* Given the widespread adaption of smartphones and corresponding digital footprints, the use of digital footprints thus has the potential to boost access to credit for some of the currently 2 billion working-age adults worldwide who lack access to services in the formal financial sector, thereby fostering financial inclusion and lowering inequality.
* Lucas critique
* Regulators: lenders using digital footprints are therefore likely to face scrutiny whether the digital footprint proxies for information violating fair lending acts (see also Fuster et al. 2018 on this issue)

# PAPER CITATI

## Blattner

* We focus on the mortgage market which plays a prominent role in the persistence of wealth gaps across generations (Charles and Hurst, 2003; Kuhn et al., 2020), with historically disadvantaged groups being less likely to transition into home ownership and build home equity (Charles and Hurst, 2002)
* Recent work has also warned that more flexible statistical technology such as machine learning can reduce overall loan approval rates for disadvantaged groups (Fuster et al., 2020), and that modern, FinTech underwriting continues to generate cross-group disparities in loan terms (Bartlett et al., 2019); credit scores likewise are seen to play a role in geographic misallocation in the US mortgage market (Hurst et al., 2016).
* persistent **policy concerns about equity** across consumers **in credit scoring** (Avery et al., 2009, 2012; Traub, 2013).

## Gambacorta

* All about fintech & BIG data
* **Frost et al (2019)**: when it comes to predicting loss rates, credit scoring techniques based on big data and machine learning have so far outperformed credit bureau ratings. A key question here is whether this outperformance will persist through a full business and financial cycle. new forms of non-prudent risk-taking???
* Another stream of the literature analyses **unfair price discrimination**. In particular, sophisticated machine learning algorithms may not be as neutral as their mathematical nature suggests at first glance. Even though artificial intelligence and machine learning algorithms are neither trained nor fed with **protected characteristics** such as race, religion, gender or disability, they are **able to triangulate such information**. Using data on US mortgages, **Fuster et al (2019)** find that Black and Hispanic borrowers are disproportionately less likely to gain from the introduction of machine learning in credit scoring models

## Singh

List of examples of companies who created committees/guidelines for fairness in ML

## The impact of credit scoring on consumer lending (einav 2013) - FINITO

Using data from the Survey of Consumer Finances, Edelberg (2006) documents the extent of this transformation. She finds that as a result the correlation between loan pricing and estimated and realized default risk has sharply increased. Grodzicki (2012) documents a similar pattern in the credit card industry and ties it specifically to lenders’ investments in information technology.

🡪 RIFORMULATO 🡪 Today, automated credit scoring has become a standard practice in pricing mortgages, auto loans, and unsecured credit. Consequently, the correlation between loan pricing and estimated and realized default risk has notably strengthened (Edelberg, 2006). Similar observations have been made by Grodzicki (2012) in the credit card industry, attributing this shift to lenders' increased investments in information technology.

## Altri paper

* Willen and Zhang, 2021 (Taken from bartlett 2022):

They point out that **econometric problems** can arise when trying to detect discrimination by regressing interest rate on race and points (or vice versa) if

1. Borrowers are choosing loans from menus with different points/rate combinations; and

2. Those menus are heterogeneous in level and/or slope across lenders.

* Possible literature for conclusion for why minorities do not actively look for better rates (from Bartlett 2022):

Woodward, S.E., 2008. A Study of Closing Costs for FHA Mortgages. Re-

port. U.S. Department of Housing and Urban Development. Woodward, S.E., Hall, R.E., 2012. Diagnosing consumer confusion and sub- -optimal shopping effort: theory and mortgage-market evidence. Am.

Econ. Rev. 102, 3249–3276.

* **FORST**: su espansione di FinTech e possibili conseguenze, utile per intro/conclusione