Chapter 6

Testing for Discrimination in Mortgage Pricing with Data Envelopment Analysis

Dominic Russel and Jason Dietrich
March 2026

I. Introduction

Mortgage pricing is characterized by complex tradeoffs between rate, points, and fees, which have different units of measure.[1] For example, depending on an applicant's time horizon, she may opt to pay discount points to buy-down the rate, or accept a higher rate in exchange for a rebate to help with closing costs. Additionally, the lender may increase origination fees to offset an applicant's desire to pay no discount points or to receive a lower rate. These complexities afford lenders many options for discriminating against applicants and create significant challenges for fair lending analysts trying to detect prohibited discrimination. Typically, fair lending analyses focus on a single pricing measure, such as interest rate (e.g., Gordon W. Crawford and Eric Rosenblatt, 1999; Frank E. Nothaft and Vanessa G. Perry, 2002; Thomas P. Boehm et al., 2006); a composite measure, such as Annual Percentage Rate (APR), which combines rate, points and fees into one measure of price (e.g., Marsha Courchane, 2007); or a lender-specific discretionary function of price, such as overages or underages (e.g., Marsha Courchane and David Nickerson, 1997; Harold A. Black et al., 2003). None of these captures all the tradeoffs made by each party in mortgage pricing, or address lenders' ability to discriminate through different pricing components or combinations of components.

This paper explores an alternative approach to analyzing disparities in mortgage pricing decisions using Data Envelopment Analysis (DEA). DEA allows the analyst to investigate potential disparities in the components of pricing *without* requiring her to make assumptions about the relationship between different components of price. Using the non-parametric, linear

[1] While points can be converted to dollars with information about loan size, we will choose to leave points in their original measure to most closely capture the tradeoffs made by borrowers and lenders in mortgage pricing. For example, mortgage points are largely chosen by the borrower, whereas other fees are largely set by the lender; therefore, borrowers, who may have information the lender does not about the length of time they will hold their mortgage, may select different point-rate tradeoffs than a lender making point-fee tradeoffs. For some transactions, there may be interchangeability between points and fees (a result we will include for some applicants in our data generating process), but this is unlikely to be the case for all lenders or all transactions.

programming approach, the analyst derives a "frontier" of applicants with the highest ("worst") pricing conditional on product and applicant characteristics. If a demographic group is disproportionately on or near the frontier relative to another, it provides evidence of discrimination. For each applicant below the frontier, DEA provides a measure indicating how much higher the pricing components were for other similar applicants — these results can be used to identify damages to applicants where a violation of law is identified.

Since the foundational work on frontiers by M.J. Farrell (1957) and the seminal work on DEA by Abraham Charnes, William W. Cooper, and Edwardo Rhodes (1978), DEA has been applied to a wide variety of settings. However, it has not previously been applied to fair lending analyses even though its central notion of "relative efficiency" is consistent with the long-standing fair lending concept of comparing outcomes of similarly-situated applicants. Perhaps the closest related study is by Bernd Ebersberger et al. (1997), which used DEA to analyze potential discrimination by landlords.

In this paper, we assess how effective DEA is at identifying patterns of discrimination and individual disadvantaged borrowers. We use simulated data based on applications analyzed during a typical supervisory fair lending examination. To benchmark how accurately DEA-based methods identify disadvantaged applicants, we compare it with standard fair lending tests. We use multiple hypothetical scenarios to test the sensitivity of our results to our data generating assumptions. We find that DEA-based approaches outperform standard APR-based regression approaches. More broadly, our results suggest DEA may be useful in a wide variety of settings to detect discrimination in which tradeoffs exist between multiple outcome measures.

I.1 Context and Related Literature

In the United States, fair lending laws restrict lenders from prohibited discrimination against minority borrowers.[2] Fair lending analysts at federal agencies conduct statistical analyses during supervisory examinations to help enforce these laws. With respect to mortgage originations, these analyses historically focused on redlining[3] and discrimination in underwriting. As lenders introduced risk-based pricing in the mid-1990s, fair lending analysts also began to assess whether pricing terms and conditions were fairly and consistently applied across borrowers. Understanding whether lenders discriminate against minority borrowers along each of these dimensions (redlining, underwriting, and pricing) is important not only academically, but for enforcing fair lending laws.[4]

In this paper we focus in particular on mortgage pricing, which has received relatively less attention in the literature given the recency with which lenders have adopted risk-based pricing, the complexity in analyzing the trade-offs between pricing components, and the limited available data on these components. The methods we introduce are intended to aid fair lending analysts attempting to detect prohibited discrimination in pricing given these complexities. (Accordingly, when we refer to "fair lending" below, our focus will be mortgage pricing in particular.) Improvements in testing accuracy are important from a consumer welfare perspective, since even small discriminatory upcharges in rate or points paid can translate into large financial impacts through interest charges and closing costs.

---

[2] For example, under the Equal Credit Opportunity Act (ECOA) and the Fair Housing Act (FHA).
[3] Redlining is "a term used for an illegal practice where people living in a certain area or neighborhood are not given the same access to credit as people in other areas or neighborhoods on the basis of race, color, or for some other prohibited reason" (Daniel Dodd-Ramirez and Patrice A. Ficklin, 2016).
[4] Stephen L. Ross and John Yinger (2002) provide a general overview of mortgage discrimination and fair lending research. Works that study mortgage underwriting include Alicia H. Munnell et al. (1996), Helen F. Ladd (1998), and Jason Dietrich (2005). Those that study mortgage redlining include Andrew Holmes and Paul Horvitz (1994) and Geoffrey M.B. Tootell (1996). More recently Andrew Hanson et al. (2016) studies differences in financial institution service quality between prospective minority and non-minority borrowers.

In addition to the fair lending works cited above, a number of more recent papers have studied the question of whether minority borrowers face mortgage pricing discrimination. For example, using pre-financial crisis data from the Survey of Consumer Finances, Ping Cheng, Zhenguo Lin, and Yingchun Liu (2015) find that, without controlling for discount points, black borrowers on average paid about 29 basis points more than comparable white borrowers. Andra C. Ghent, Rubén Hernández-Murillo, and Michael T. Owyang (2014) find similar results using loan-level data, again without information about discount points. Robert Bartlett et al. (2019) study the post-crisis mortgage market, finding differences in interest rates between similar minority and non-minority borrowers, but also do not directly observe points or fees. Neil Bhutta and Aurel Hizmo (2021) study a similar era, but do observe these components of price and use them to show that minority and white borrowers make different points-rate tradeoffs. They find that, conditional on covariates, when the two pay the same interest rate for FHA loans they do not pay different amounts in fees. Perhaps most closely related to our paper is David Zhang and Paul Willen (2020), who frame the pricing tradeoffs challenge as a "menu problem" and illustrate why, for example, the approach taken by Neil Bhutta and Aurel Hizmo (2021) could lead to contradictory estimates depending on whether rates or points are held constant. Using a new "difference in menus" metric, they find differences in mortgage pricing by race for conforming mortgages, but not FHA mortgages. As opposed to the methods they introduce, our approach utilizes a set of well-developed tools already widely used in operations research.[5]

---

[5] Another key difference in our approaches is that our DEA method focuses on differences in *outcomes* rather than differences in *menus* that borrowers select from. While we believe in practical fair lending applications our approach will capture the object of interest (whether lenders are discriminating against minority borrowers), David Zhang and Paul Willen (2020) present evidence that differences between groups along these two dimensions (outcomes and menus) may not always be the same.

Our paper makes three primary contributions to the literature. First, using information from past fair lending supervisory work we are able to construct a detailed model of the general decision-making process lenders use when pricing mortgages. We use this model to inform our simulation-based data generating process. We are not aware of any other studies that have developed such a detailed model. Second, we build upon the recent studies that focus on discrimination in mortgage pricing in the context of pricing component tradeoffs. Because these components afford lenders many options for discriminating against borrowers, developing a better understanding of these tradeoffs is particularly important for identifying discrimination and estimating the magnitude of discrimination. Finally, our application of DEA to mortgage pricing discrimination introduces both a new set of tools to the literature focused on mortgage pricing discrimination and a new empirical application to the operations research literature.

The remainder of the paper is structured as follows. Section II presents the DEA approach. Section III summarizes the data and the Data Generating Process (DGP) used to construct our loan-level data. Section IV presents the results, including a comparison to a single composite measure (APR) regression approach. Section V concludes the discussion.

II. Empirical Approach: Data Envelopment Analysis

DEA is a non-parametric, linear programming approach, grounded in microeconomic production theory. A firm's production function expresses the relationship between inputs employed and outputs produced. The production possibility frontier (PPF) shows the maximum set of outputs possible given a set of inputs, assuming all inputs are used efficiently. DEA uses data on decision-making units (DMUs) to empirically identify the PPF. The input/output set of each DMU is compared to the PPF to construct relative "efficiency" measures.

A key insight of this paper is that by letting: (a) each pricing decision be the DMUs, (b) legitimate determinants of price (product characteristics and applicants' risk profiles) be the inputs, and (c) components of price (rate, points, and fees) be the outputs, the empirically derived PPF reflects the applicants with the highest pricing conditional on legitimate determinants of price. Therefore, for a given demographic group, higher DEA "efficiency" – henceforth referred to as "pricing intensity" – would suggest evidence of a pattern of discrimination.

Formally, suppose we have n loan pricing decisions which take an m-dimensional input vector (applicant and product characteristics) to select a 3-dimensional output vector (rate, discount points, and fees). DEA allows us to measure the pricing intensity of each loan (i.e., how favorable the pricing appears from the lender's perspective) in comparison to the other loans. For pricing decision j, we define the following objective function:[6]

$$PricingIntensity_j = \max_{\delta_{pj},\gamma_{rj}} \frac{\sum_{r=1}^{3} \gamma_{rj} y_{rj}}{\sum_{p=1}^{m} \delta_{pj} x_{pj} + \delta_{0j}} \quad (1)$$

subject to the constraints:

$$\frac{\sum_{r=1}^{3} \gamma_{rj} y_{ri}}{\sum_{p=1}^{m} \delta_{pj} x_{pi} + \delta_{0j}} \leq 1 \quad (2)$$

$$\gamma_{rj} \geq 0; \delta_{pj} \geq 0; \ \gamma_{0j} \text{ unconstrained} \quad (3)$$

$\forall$ $r = 1, 2, 3$ outputs;

$p = 1, \ldots, m$ inputs;

$i, j = 1, \ldots, n$ pricing decisions

For pricing decision j we are finding m input weights δ_{pj} and three output weights γ_{rj}, given m input applicant/product characteristics x_{pj} and three output pricing components y_{rj}. The free

[6] When we run our DEA, we solve the more tractable dual of an equivalent linear programming problem, as described by Rajiv D. Banker, Abraham Charnes, William W. Cooper (1984). We present the more intuitive fractional version here.

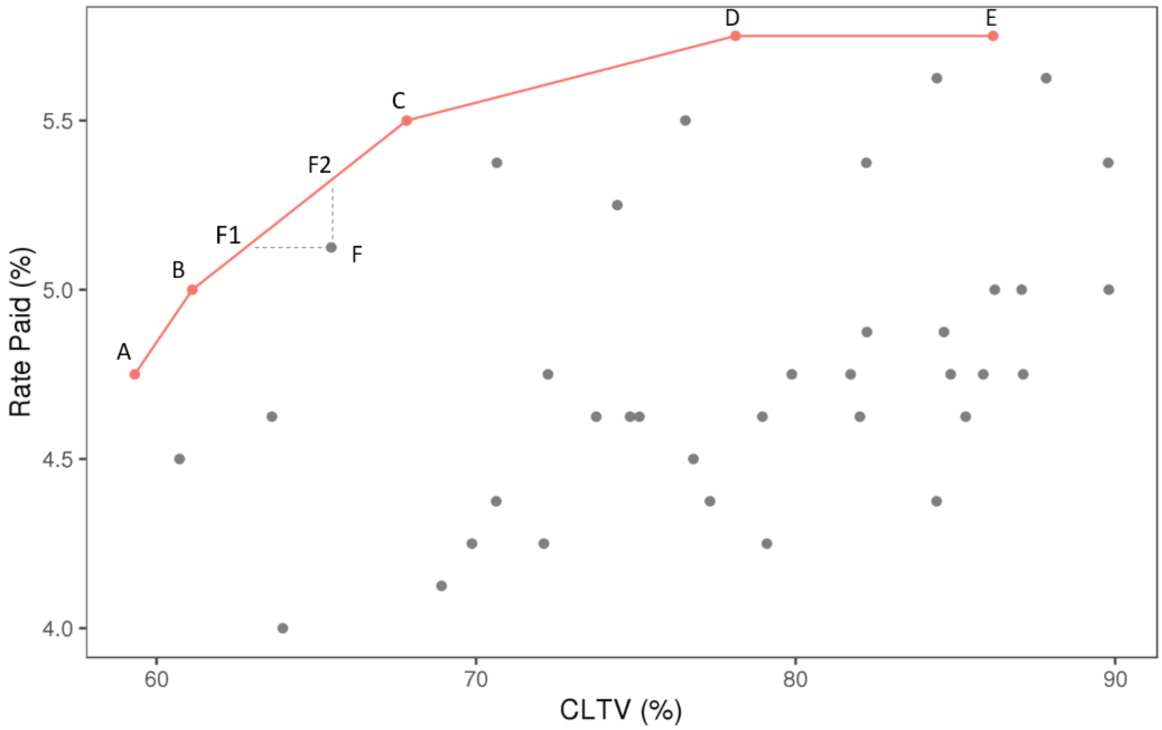
parameter δ_0 allows for variable returns to scale.[7] The max found in equation (1) is the pricing intensity (efficiency) ratio, our primary metric of interest. Intuitively, larger outputs and smaller inputs lead to higher pricing intensity ratios. The first constraint (2) guarantees that the pricing intensity ratio of any decision $(i = 1, \dots, n)$ using the same input and output weights used here for pricing decision j are never greater than one. Values of one represent maximally intense pricing decisions and values between 0 and 1 represent less intense pricing.[8] For example, suppose pricing decisions A and B are identical on inputs, points, and fees, but B has a higher rate. A will necessarily be off the frontier because B will have a higher pricing intensity using A's weights and B cannot have a pricing intensity greater than one.

To provide some additional intuition, we use a one-input, one-output pricing example where combined loan-to-value (CLTV) is the only input and interest rate is the only output. Figure 1 shows a scatter plot of CLTV and interest rate for 40 originated mortgages. The DEA frontier is defined by the concave boundary created by five loans A, B, C, D, and E, and shows the highest possible interest rates available across all possible CLTV values. The concavity of ABCDE is a result of constraint (2) and is indicative of variable returns to scale (while a linear frontier would suggest constant returns to scale).

This frontier shows the maximum rate possible for a given CLTV. The pricing intensity of each loan is determined by its distance from the frontier. All five loans on the frontier are maximally priced, in the sense that a higher rate could not be obtained given the CLTV value, and would have a pricing intensity of 1. Each loan below the frontier has a pricing intensity that decreases with distance from the frontier (see Loan F for example). For fair lending purposes where

[7] The model without this free parameter is the original constant returns to scale DEA model, introduced by Charnes, Cooper, and Rhodes (1978) widely known as the CCR model. The model presented here is referred to as the variable returns to scale or the Banker, Charnes, and Cooper (BCC) model, named for the authors who advanced it (1984).
[8] A value of 1 represents "weak efficiency" here as this representation does not incorporate slacks.

Figure 1: Example of Data Envelopment Analysis



Note: Figure shows a one-input (CLTV) and one-output (interest rate paid) DEA example. Each point is a mortgage pricing decision. The DEA frontier is defined by the concave boundary created by five loans A, B, C, D, and E. F is a loan below the frontier.

applicant/product characteristics are fixed and pricing is chosen, we measure pricing intensity as $Rate_F/Rate_{F2}$. This represents how much higher the rate could have been given the CLTV. (Later, we will extend this to multiple dimensions.) From the lender's perspective, loans on or close to the frontier are preferable, as they have the highest rate possible conditional on CLTV. From an applicant's perspective, points further from the frontier are preferable. Higher price intensity for minorities[9] on average would suggest fair lending risk.[10]

[9] Throughout, we use the term minorities to refer generally to historically disadvantaged minority racial groups. In practice, this group could be any potentially disadvantaged group.

[10] One final concept of relevance in DEA is slack. For applicants with segment DE as a reference, inefficiency in an output-oriented approach would be measured as a vertical distance from that applicant to the segment. Given that this segment is horizontal, all applicants on this segment would be considered on the frontier and efficient. However, for any point on that segment, input (CLTV) could have been higher with no change in output (rate). This adjustment is called slack and captures possible adjustments to inputs (outputs) with no corresponding change to any outputs (inputs). Slacks do not impact the optimal q0 from equations (1) – (3), and can therefore be determined through a second, independent optimization process as described by Meilin Wen (2015).

To show how our maximization problem leads to this frontier in a single-input, single-output example, we can re-write it geometrically. For each pricing decision j (a rate/CLTV "point") we hold $\gamma_{1,j}$ fixed at one and seek to find $\delta_{1,j}$ and $\delta_{0,j}$ that will define some line $Rate = \delta_{1,j} * CLTV + \delta_{0,j}$. This line must have non-negative slope ($\delta_{1,j} \geq 0$; constraint 3), run through or above all rate/CLTV points ($\delta_{1,j} * CLTV_i + \delta_{0,j} - Rate_i \geq 0, \forall i$; constraint 2), and have minimum vertical distance from j ($\min_{\delta_{1,j},\delta_{0,j}} [\delta_{1,j} * CLTV_j + \delta_{0,j} - Rate_j]$; equation 1). A loan will be on the frontier if a line (generally non-unique) runs through its rate/CLTV point and satisfies the two geometric constraints. For pricing decisions off the frontier (e.g., F), the line that minimizes geometric equation (1) and satisfies the constraints is the line through the rate/CTLV points on the frontier with nearest smaller and larger CLTVs (e.g., B and C). These latter segments form the frontier. Using $\delta_{1,j}, \delta_{0,j}$ we calculate pricing intensity as in equation 1: $PricingIntensity_j = \frac{Rate_j}{\delta_{1,j}*CLTV_j+\delta_{0,j}}$, now a measure of the vertical distance between the rate and the frontier. While this simple example includes only one input and one output, we next extend our original maximization problem to a more realistic application with three outputs (rate, discount points, and fees) and multiple inputs (FICO score, CLTV, and loan amount).

III. Data Generating Process

To assess how effectively DEA detects patterns of discrimination, we construct simulated datasets where we control who is disadvantaged and by how much. We begin with data consisting of product and applicant characteristics based on applications analyzed during a typical supervisory fair lending exam. We generate a sample of 2,569 loans that are owner-occupied, 1-4 family, 30-year fixed rate, and conventional mortgages in the Retail channel with a 30-day commitment period. We add simplifying assumptions that applicants are from the same state and that pricing does not change over time. Table 1 presents summary statistics for three product and applicant characteristics, by minority status, for the sample. On average, minorities have lower loan amounts, higher CLTVs, and lower FICO scores.

Table 1: Summary Statistics by Minority Status

Minority	Variable	# of loans	Mean	Std. Deviation
Y	Loan Amount	370	$224,370.27	$125,321.14
Y	CLTV	370	71.77	17.88
Y	FICO	370	730.02	61.35
N	Loan Amount	2,199	$247,068.67	$137,551.45
N	CLTV	2,199	69.52	17.86
N	FICO	2,199	754.24	52.51

Note: Table presents summary statistics describing the sample of mortgages. Statistics are presented separately for minority and non-minority borrowers.

We next assign rate, points, and fees, simulating the policies of a Retail lender that uses risk-based pricing.[11] We describe this method of pricing in Appendix A. Equations A1-A6 formalize our description and the foundation of our DGP. Appendix B provides the rate sheet with the rate/point schedule and the Loan-Level Price Adjustments (LLPAs) consisting of loan amount, CLTV, and FICO score. The LLPAs are based on the standard LLPA grid Fannie Mae publishes.[12] With our simplifying assumptions, each applicant faced the same rate/point schedule prior to the LLPA adjustments, so the only product and applicant characteristics that impacted pricing will be loan amount, CLTV, and FICO score.[13]

[11] Wholesale and Correspondent lending create an additional layer of complexity with additional decision-makers. See Amany El Anshasy et al. (2004), Stanley D. Longhofer and Paul S. Calem (1999), and Morris M. Kleiner and Richard M. Todd (2009) for more information on analyzing mortgage pricing by brokers.
[12] Fannie Mae's LLPA matrix is available at http://www.fanniemae.com/content/pricing/llpa-matrix.pdf.
[13] Mortgage pricing is more complex then represented here. For example, lenders typically consider additional factors when pricing mortgage loans, different lenders consider different factors, and factors differ by product, program, and

Once the total LLPAs are calculated for each applicant, we construct applicant specific, post-LLPA rate/point schedules. We draw from a normal distribution to determine the specific rate, and corresponding discount points, for each applicant. During fair lending examinations, we typically find non-minority applicants are often less liquidity constrained and therefore more likely to pay points to buy down the rate.[14] To incorporate this finding into our data, the normal distribution is centered at the post-LLPA par rate for minorities and at the post-LLPA par rate less 25 basis points (bps) for non-minorities.[15] The variances of the normal distribution we draw from is set to 5 basis points for both groups.

The last two pricing components to be determined are origination points and third-party fees. Using a common industry practice, we set origination points to 1 for all loans.[16] Fees are constructed by drawing from a normal distribution with mean $5,000 and standard deviation $850.[17] For simplification, we combine origination points, as well as third party fees that impact the APR, into an aggregate "fees" measure, which is one of the three pricing outcomes analyzed.

We now have three components of price (rate, discount points, and fees), which incorporate the explicit tradeoffs between rate and points from the rate sheet. During actual transactions, the lender can obtain its total reservation price through any of the three pricing components. To reflect this, we extend the potential tradeoffs between rate and points to also include fee tradeoffs. We assume that 30 percent of applicants (selected from a uniform distribution) express preferences about specific pricing components. For these applicants, discount points are randomly adjusted,

time. A real-world fair lending analysis would focus on a particular lender/product and could incorporate additional factors as inputs.

[14] This is consistent with racial wealth disparities in the U.S. (e.g. Neil Bhutta et al., 2020).

[15] A 0-discount point rate is not available on the LLPA-adjusted rate/point schedule for every applicant. In such instances, we use as the post-LLPA rate corresponding to the discount points closest to 0.

[16] See Rebecca Lake (2019). This is also consistent with application data in a typical fair lending exam.

[17] In total, this combination of origination points and other fees imposes closing costs in line with estimate from Deborah Kearns and Barbara Marquand (2019), on the lower end of between 2 and 5 percent of the property cost.

and then an exact offset adjustment is applied to fees so that the lender obtains the same reservation price.[18] The remaining 70 percent of applicants express no preferences about the components of price, so pricing is based on the lender's stated policies. At this point, we have a sample of 2,569 loans with rate, discount points, and fees.

The last step of the pricing process is to incorporate disadvantage to minority applicants. We chose to focus on a scenario with a discriminatory loan officer(s), who opportunistically makes discretionary pricing adjustments through one of the three pricing channels to disadvantage minority borrowers whenever possible. We do this with the understanding that there are an unlimited number of possible discriminatory scenarios, and that all results may be dependent on our simulation choices. To test the sensitivity of our results, we generate three different discriminatory scenarios. In each scenario, only minority borrowers are disadvantaged. The first scenario (no disadvantage) is based solely on the DGP in equations A1-A6 in Appendix A and described above. The second is a moderate disadvantage scenario: among minorities, 40 percent receive no disadvantage, 20 percent receive disadvantage through rate, 20 percent receive disadvantage through discount points, and 20 percent receive disadvantage through fees (the sum of origination points and APR-impacting fees). No applicant is disadvantaged through multiple pricing components. The amount of disadvantage is determined by random draws from a chi-square distribution with a mean of 2 points and standard deviation of 0.7. The third is an extreme disadvantage scenario: among minorities, 33 percent receive disadvantage through rate, 33 percent receive disadvantage through discount points, and 33 percent receive disadvantage through fees.

[18] Borrowers typically negotiate with lenders over the number of discount points they must pay or the number of rebate points they would receive for a given interest rate, but do not negotiate over fees charged. Lenders use fee adjustments to offset concessions to borrowers with preferences over paying no points or fewer points than specified by the relevant rate/point tradeoff. The adjustments here capture those pricing decisions. Whether applicants express preferences for paying discount points, origination points or fees is not correlated with race.

No applicant is disadvantaged through multiple pricing components. The amount of disadvantage is determined by random draws from a chi-square distribution with a mean of 4 points and standard deviation of 1.50. For the second and third scenarios, when disadvantage enters via rate, the disadvantage amount from the random draw is divided by 4 to reflect the general relationship of 0.25 bps in rate equaling 1 point. When disadvantage enters via fees, the disadvantage amount from the random draw is converted into dollars.

At this point, the DGP is complete: rate, points and fees have been determined for each of the 2,569 applicants. Differences in rates, points, and fees across applicants are due to three factors: 1) differences in creditworthiness as captured by the LLPAs (FICO score, CLTV, and loan amount); 2) differences in applicants' choices to buy-down or buy-up rate; and 3) discriminatory premiums. Typically, differences in rate, points, and fees will also be driven by additional factors such as property type, occupancy, product, program, timing, commitment period, geography, and channel that lenders use to identify the specific rate sheet to apply and the specific rate/point schedule to use. To simplify this study, we have held these factors constant for all applicants. Except for timing, all of these factors are categorical variables, so in real-world applications, separate DEAs would be conducted for each value of these categorical variables.

Table 2 presents summary statistics for the three scenarios. In the "no disadvantage" scenario, minorities paid 28 bps more on average in rate, 2.00 points less on average in discount points, $89.65 less on average in fees, and 14 bps more on average in APR. These differences are driven by product and applicant characteristics, as well as differences in each group's preference for buying down (or up) the rate. This highlights a challenge fair lending analysts must address when testing for discrimination: tradeoffs between the different components of pricing and borrower point/rate choices, resulted in fairly sizeable disparities in each of the pricing components even though the DGP contained no discriminatory premium. In the moderate disadvantage scenario, the simulated data shows that minorities paid 37 bps more on average in rate, 1.65 points less on average in discount points, $898.98 more on average in fees, and 30 bps more on average

Table 2: Summary Statistics of Pricing Measures by Disadvantage and Minority Status

No Discrimination DGP

Minority	Variable	# of loans	Mean	Std. Deviation
Y	Rate	370	4.83	0.26
Y	Discount Points	370	0.22	1.92
Y	Fees	370	$5,742.65	$1,811.91
Y	APR	370	5.13	0.21
Y	LLPAs	370	0.54	0.87
N	Rate	2,199	4.55	0.23
N	Discount Points	2,199	2.22	1.87
N	Fees	2,199	$5,832.30	$2,033.22
N	APR	2,199	4.99	0.18
N	LLPAs	2,199	0.27	0.72

Moderate Discrimination DGP

Minority	Variable	# of loans	Mean	Std. Deviation
Y	Rate	370	4.92	0.32
Y	Discount Points	370	0.57	2.04
Y	Fees	370	$6,731.28	$3,274.43
Y	APR	370	5.29	0.29
Y	LLPAs	370	0.54	0.87
N	Rate	2,199	4.55	0.23
N	Discount Points	2,199	2.22	1.87
N	Fees	2,199	$5,832.30	$2,033.22
N	APR	2,199	4.99	0.18
N	LLPAs	2,199	0.27	0.72

Extreme Discrimination DGP

Minority	Variable	# of loans	Mean	Std. Deviation
Y	Rate	370	5.22	0.71
Y	Discount Points	370	1.73	3.14
Y	Fees	370	$9,852.62	$8,088.46
Y	APR	370	5.81	0.54
Y	LLPAs	370	0.54	0.87
N	Rate	2,199	4.55	0.23
N	Discount Points	2,199	2.22	1.87
N	Fees	2,199	$5,832.30	$2,033.22
N	APR	2,199	4.99	0.18
N	LLPAs	2,199	0.27	0.72

Note: Table presents summary statistics describing pricing of the sample of mortgages. Statistics are presented separately for minority and non-minority borrowers, and by discrimination scenario.

in APR. Finally, in the extreme disadvantage scenario, the simulated data shows that minorities paid 67 bps more on average in rate, 0.49 points less in discount points, $4,020.32 more in fees, and 82 bps more in APR.

The differences for the moderate and extreme disadvantage scenarios are driven by differences in product and applicant characteristics, differences in each group's preference for buying down or buying up the rate, and the disadvantage we injected into the data. Table 3 presents additional summary information about the number of minorities that were disadvantaged and the amount of disadvantage for each pricing component.[19] The goal of the DEA presented below is to detect these patterns of discrimination and identify the specific minorities that were disadvantaged.

[19] Non-minorities are never disadvantaged in any of the DGPs, so Tables 3 only presents results for minorities. No minorities are disadvantaged in the "No Disadvantage" scenario, so Table 3 only presents results for the "Moderate" and "Extreme" disadvantage scenarios.

Table 3: Simulated Disadvantage in Rate, Discount Points and Fees for Minorities

Moderate Discrimination DGP

	# Disadvantaged	Average disadvantage for disadvantaged minorities	Average disadvantage for all minorities
Rate	76 of 370	45.33 bps	9.31 bps
Points	68 of 370	1.92 pts	0.35 pts
Fees	80 of 370	$4,572.39	$988.62

Extreme Discrimination DGP

	# Disadvantaged	Average disadvantage for disadvantaged minorities	Average disadvantage for all minorities
Rate	120 of 370	121.48 bps	39.40 bps
Points	117 of 370	4.78 pts	1.51 pts
Fees	133 of 370	$11,433.74	$4,109.97

Note: Table presents summary statistics describing the disadvantage added to minority borrowers' pricing. Summaries are presented separately for the moderate and extreme discrimination scenarios.

Before presenting the results of the DEA, one final set of data transformations is necessary. Traditionally, DEA is used in settings where positive amounts of continuous measures of input are used to produce positive amounts of continuous measures of output, typically at a non-increasing rate. Data for fair lending exams are somewhat different with non-positive values for some variables (for example discount points); non-continuous variables (for example property type); and inverse relationships between some inputs (for example FICO score) and pricing outputs. As a result, prior to running the DEA, a number of data transformations are needed. First, following Joesph Sarkis (2007) and Jesus T. Pastor and Jose L. Ruiz (2007) a constant is added to outputs with negative values to make all values greater than 0 (specifically, we add four to discount points). Second, following Herbert F. Lewis and Thomas R. Sexton (2004), for FICO score, which has an inverse relationship with pricing, we use (901 – FICO) in the DEA instead of FICO. Third, following Gerrit Lober and Matthias Staat (2010) and Banker and Morey (1986), to incorporate categorical variables into the DEA, we have limited our data to a homogenous subset of originations for 30-year, conventional, conforming, fixed rate Retail loans with the same commitment period, occupancy status, and property type and in the same state. For pricing analyses that include categorical inputs, a fair lending analyst could run separate DEAs for the different values. Finally, because DEA compares each applicant to the "best" applicant, it can be very sensitive to outliers (R.G. Dyson et.al, 2001; Rayna Brown, 2006). Whereas our DGP ensures no outliers, a simple outlier analysis could be run for exams.

IV. Results

Fair lending analyses have three primary objectives: 1) identifying patterns of discrimination, 2) identifying disadvantaged consumers, and 3) estimating damages for disadvantaged consumers for the purpose of compensation. In this paper, we focus on the first two objectives: identifying patterns of discrimination and identifying disadvantaged consumers. For each of these objectives, we conduct analyses using DEA on the simulated data as if it were an actual exam.[20] We then assess how well DEA-based methods perform in comparison to an APR-based regression. Objective three (estimating the necessary compensation for damages) requires the fair lending analyst to make assumptions about the relationship between the different components of pricing. Because the primary advantage of the DEA technique is the lack of such assumptions, we do not address this topic in this paper.

[20] During fair lending exams, analysts have access to lender's policies and procedures and therefore know all factors the lender considers when making decisions on applications for credit. However, analysts do not always have data for every factor a lender considers, so omitted variable bias is a common challenge. Similar to fair lending exams, we know all factors that affect pricing outcomes, but unlike many exams, we have electronic data for each factor as well. Assessing the impact of omitted variables on the effectiveness of the DEA at detecting discrimination is an important question, but beyond the scope of this paper.

Patterns of discrimination

To begin assessing how well DEA identifies patterns of disadvantage, we compare the average pricing intensity scores for minorities and non-minorities, as well as the share of minorities and non-minorities on the pricing frontier. Table 4 presents these results for all three disadvantage scenarios.

Table 4: Estimated Patterns of Discrimination – Data Envelopment Analysis

	No Disadvantage Scenario	Moderate Disadvantage Scenario	Extreme Disadvantage Scenario
All Applicants			
Average pricing intensity	0.965	0.923	0.822
Percentage on Frontier	4.75% (122 of 2,569)	3.85% (99 of 2,569)	3.74% (96 of 2,569)
Minorities			
Average pricing intensity	0.965	0.938	0.912
Percentage on Frontier	7.30% (27 of 370)	11.35% (42 of 370)	17.03% (63 of 370)
Non-Minorities			
Average pricing intensity	0.965	0.920	0.807
Percentage on Frontier	4.32% (95 of 2,199)	2.59% (57 of 2,199)	1.50% (33 of 2,199)
Average Pricing Intensity Disparity			
Minority – Non-minority disparity	0.01 pps	1.82 pps	10.52 pps
Two proportion t-test p value	0.926	<0.0001	<0.0001

Note: Table presents the average DEA pricing intensity and the precent of borrowers on the frontier separately for all borrowers, minority borrowers, and non-minority borrowers. The bottom rows present the average pricing intensity disparity between minority and non-minority borrowers, and the p-value from a two-proportion t-test for equivalence. Column 1 presents measures from the no discrimination scenario, column 2 from the moderate discrimination scenario, and column 3 from the extreme discrimination scenario.

Consistent with the DGP, the results show that the percentage of minorities on the frontier increases from the no disadvantage scenario to the moderate disadvantage scenario, and again to the extreme disadvantage scenario. However, even in the no disadvantage scenario, the share of minorities on the frontier is higher than the share of non-minorities. In addition, a number of non-minorities fall on the frontier. A reason for this second result is that some non-minority borrowers have no similarly-situated minority borrowers. One feature of DEA is that the results are driven entirely by comparisons between DMUs in the data. As a result, when non-minority borrowers have no minority borrowers with similar input characteristics, their pricing intensities are measured relative only to other non-minorities. As one measure of this, for a given borrower off the frontier, DEA allows us to identify a "reference set" of borrowers – those with similar inputs and higher pricing, who determine the pricing intensity for the given borrower:[21] in the moderate and extreme disadvantage scenarios, respectively, 65 percent and 73 percent of non-minorities on the frontier are only in the "reference set" of non-minorities.

The average pricing intensity score in the no disadvantage scenario is similar for minorities and non-minorities (0.9647 vs 0.9646). These pricing intensity scores suggest that, on average, components of pricing for minorities and non-minorities were 3.53 percentage points (1 - 0.9647) and 3.54 percentage points (1 - 0.9646) lower, respectively, than pricing observed for other similar-situated applicants. As one measure of patterns of discrimination, a fair lending analyst could conduct a simple t-test comparing the average pricing intensity of minorities and non-minorities. The disparity in average scores (0.01 percentage points) is statistically insignificant even at a 10 percent level.

[21] For example, in Figure 1 the references for F are B and C. Mathematically, this reference set is most easily interpreted through the dual of a linear programming problem, equivalent to the fractional maximization problem we present in equation 1.

Moving to the moderate disadvantage scenario, average pricing intensity was higher for minorities than non-minorities (0.9379 vs 0.9197). Using the same test, the disparity in average pricing intensity scores (1.82 percentage points) is statistically significant at a 1 percent level. The results for the extreme disadvantage scenario are even stronger, as minorities had much higher average pricing intensity than non-minorities (0.912 vs 0.807), a 10.52 percentage point disparity.

To benchmark how well DEA performed at identifying patterns of discrimination, we next attempt to identify discrimination using an APR-based regression method.[22] Table 5 presents the results of a regression to predict APR for all three disadvantage scenarios. Consistent with the DGP, the table shows that the coefficient estimate for minority increases from the no discrimination scenario to the moderate disadvantage scenario, and again to the extreme

Table 5: Estimated Patterns of Discrimination – APR-based regression

	Outcome: APR		
	No Discrimination Scenario	**Moderate Discrimination Scenario**	**Extreme Discrimination Scenario**
Loan Amount	0.000***	0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
CLTV	0.000***	0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
FICO	0.000***	0.001**	-0.002***
	(0.000)	(0.000)	(0.000)
Minority	0.086***	0.244***	0.762***
	(0.007)	(0.008)	(0.013)

Note: Table presents results from linear regressions on APR. The independent variables are loan amount, CLTV, FICO, and a dummy for whether the borrower is a minority. Column 1 presents results from the no discrimination scenario, column 2 from the moderate discrimination scenario, and column 3 from the extreme discrimination scenario. Significance levels: ***p < 0.01, **p < 0.05, *p < 0.10

[22] For this analysis we use the data before the transformations required for DEA described in the previous section (e.g. 901-FICO).

disadvantage scenario. However, unlike in DEA, in all three scenarios, and most importantly the no discrimination scenario, minority status is a significant factor in predicting the final APR. In the next section we will compare this APR-based regression's ability to identify true patterns of discrimination with DEA-based methods in more detail.

Identifying disadvantaged consumers

A second objective of fair lending analyses is to identify individual consumers who were disadvantaged. Table 6 explores how well DEA achieves this goal. Ideally, all disadvantaged minorities would be on the frontier with pricing intensities of one, and all other applicants would be below the frontier. The average pricing intensity for non-disadvantaged minorities in the moderate disadvantage scenario (0.9037) is lower than the average intensity for borrowers disadvantaged by rate, points, or fees (0.9705, 0.9522, and 0.9573). However, relatively few disadvantaged minorities were on the frontier, ranging from 10.29 percent to 25.00 percent depending on the source of disadvantage and the scenario. A primary reason for this is that other similarly-situated minorities received *even more* disadvantage. In the moderate disadvantage scenario, all 184 disadvantaged minorities not on the frontier referenced at least one disadvantaged minority and 111 referenced only disadvantaged minorities. In the extreme disadvantage scenario, all 307 disadvantaged minorities not on the frontier referenced at least one disadvantaged minority and 285 referenced only disadvantaged minorities.

Many researchers have sought to improve the differential capabilities of DEA beyond simple efficiency (here, pricing intensity) scores. To better identify disadvantaged borrowers, we use three methods: one based on "peeling" away frontiers, one based on a DEA intensity score cutoff, and one based on a DEA cross-intensity cutoff. Each approach allows us to further refine our categorization of pricing decisions. We then compare these methods with a simple classifier built on our APR-based regression.

Table 6: Results for Minorities Using Data Envelopment Analysis

Moderate Disadvantage Scenario

Applicants	Average disadvantage for disadvantaged minorities	# and Percent on Frontier	Average Pricing Intensity
76 Minorities disadvantaged through rate	45.33 bps	19 (25.00%)	0.9705
68 Minorities disadvantaged through points	1.92 pts	7 (10.29%)	0.9522
80 Minorities disadvantaged through fees	$4,572.39	14 (17.50%)	0.9573
146 Minorities not disadvantaged	NA	2 (1.37%)	0.9037

Extreme Disadvantage Scenario

Applicants	Average disadvantage for disadvantaged minorities	# and Percent on Frontier	Average Pricing Intensity
120 Minorities disadvantaged through rate	121.48 bps	25 (20.83%)	0.9253
117 Minorities disadvantaged through points	4.78 pts	19 (16.24%)	0.9085
133 Minorities disadvantaged through fees	$11,433.74	19 (14.29%)	0.9035
0 Minorities not disadvantaged	NA	NA	NA

Note: Table presents results from DEA for minority borrowers, separately for the moderate and extreme disadvantage scenarios. Each row includes borrowers who were disadvantaged through a particular component of price.

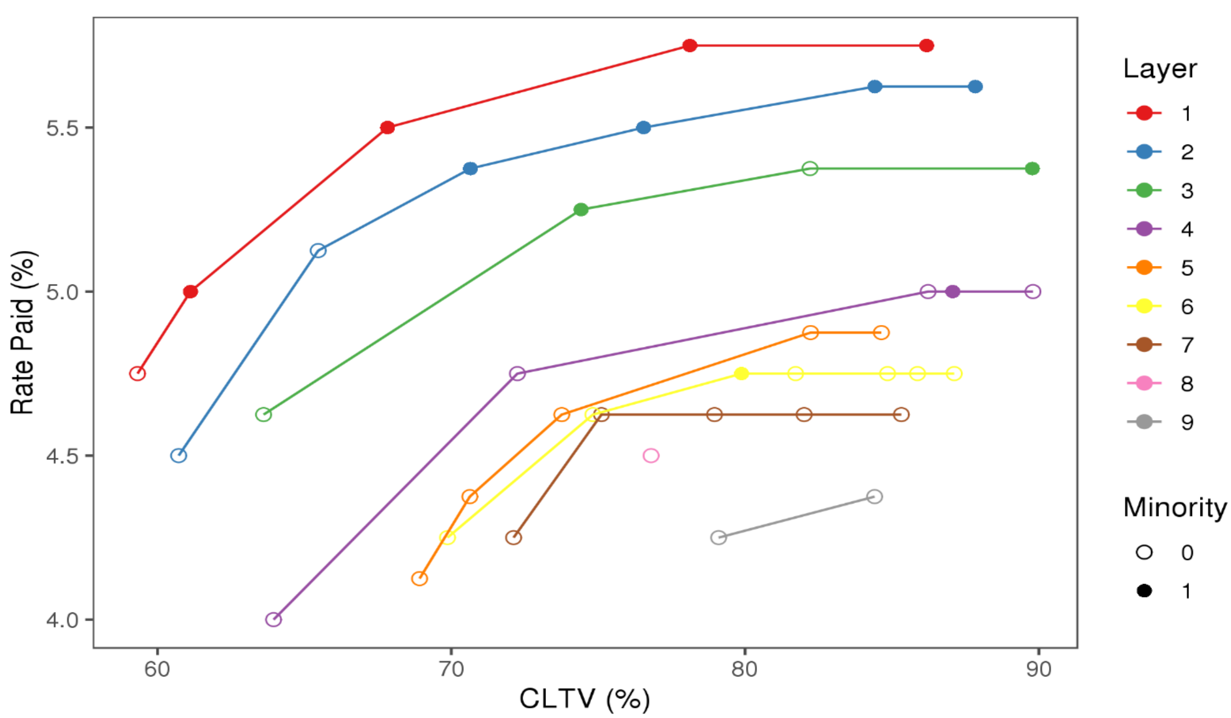
i) Peeling classifier

Our first method is built on the idea of nested DEA frontiers or "peels." Richard S. Barr, Matthew L. Durchholz, and Lawrence Seiford (2000) introduced the concept of "peeling" DEA layers to establish a new tier-based efficiency metric that could better explain the inefficiency of decision units. The approach is straightforward: DEA is first applied to the entire dataset and DMUs on the frontier are classified as members of a first layer. These DMUs are then removed from the dataset and DEA is applied to the remaining data (so long as the dataset is not empty).

The DMUs on the frontier in this second DEA are classified as members of a second layer, then removed from the data. This process continues until no data remains. With this iterative peeling approach, a disadvantaged consumer who the DEA did not originally identify as disadvantaged because of another similarly-situated consumer with worse disadvantage, would now be identified as disadvantaged once the similarly-situated consumer was excluded from the analysis.

The technique has been used in a variety of contexts including ranking performance in higher education (Marie-Laure Bougnol and José H. Dulá, 2006) and predicting bank failures prior to a major crisis (Necmi Avkiran and Lin Cai, 2014). To build intuition, we re-introduce our one-input / one-output pricing example. Figure 2 shows the subsequent concave boundaries created by removing points from the boundary and re-running DEA each time. Pricing in outer layers would

Figure 2: Example of Data Envelopment Analysis Peels



Note: Figure shows a one-input (CLTV) and one-output (interest rate paid) DEA peeling example. Each point is a mortgage pricing decision. Empty circles are loans corresponding to minority applicants, filled circles are loans corresponding to non-minority applicants. Each color line a is DEA frontier created after "peeling" the previous layer.

be categorized as relatively “more intense” than pricing in the inner layers. Disadvantaged minorities that were not identified in the first frontier because they received less disadvantage than other similarly-situated minorities would be identified in subsequent frontiers.

Returning to our three scenarios, Table 7 shows the DEA peeling results. Optimally, disadvantaged minorities would be in outer layers and other applicants in inner layers. The first two rows show the number of applicants in each layer and the share that are minorities. Consistent with the DGP, the share of minorities in the first four layers is lowest in the no disadvantage scenario, higher in the moderate disadvantage scenario, and highest in the extreme disadvantaged scenario. The share of minorities also generally decreases from the outer layers to the inner layers in the two discrimination scenarios.[23] However, the share of minorities also slightly decreases in layers 1-4 of the no discrimination scenario as well. Additionally, while the share of minorities in outer layers is greater than the share in the overall population, there remain a number of non-minorities in outer layers. As discussed above, this could occur because of non-minority borrowers without similarly-situated disadvantaged minority borrowers. In the moderate disadvantage scenario (the only scenario including both disadvantaged and non-disadvantaged minorities), more than 95 percent of minorities in the first two layers were actually disadvantaged and more than 75 percent of minorities in layers three and four were actually disadvantaged. By contrast, less than 15 percent of minorities in layers seven through twelve were actually disadvantaged.

During actual fair lending analyses, the share of truly disadvantaged minorities identified with each peel as shown in the third rows of Table 7 would not be known. So, an objective stopping criteria is needed. For some dataset with $K \geq 2$ layers and $L < K$, we can establish a binary classifier by predicting minority applicants in layers 1 through L-1 as disadvantaged and minority

[23] One way we can further explore this is a simple one-way analysis of variance F-test, to test if the proportions are equal across the layers. We find that the proportions are different and statistically significant at the 1-percent level.

Table 7: Results from Data Envelopment Analysis Peeling

No Discrimination Scenario

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10	Layer 11
Number of applicants on the frontier	122	225	325	361	339	343	331	252	181	86	4
Number & percent minority	27 (22.1%)	42 (18.7%)	48 (14.8%)	48 (13.3%)	44 (13.0%)	56 (16.3%)	46 (13.9%)	40 (15.9%)	13 (7.2%)	6 (7.0%)	0 (0.0%)
Number & percent of minorities disadvantaged	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Moderate Disadvantage Scenario

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10	Layer 11	Layer 12
Number of applicants on the frontier	99	175	241	304	319	330	329	247	224	172	101	28
Number & percent minority	42 (42.4%)	60 (34.3%)	63 (26.1%)	58 (19.1%)	33 (10.3%)	34 (10.3%)	28 (8.5%)	26 (10.5%)	10 (4.5%)	13 (7.6%)	1 (1.0%)	2 (7.1%)
Number & percent of minorities disadvantaged	40 (95.2%)	57 (95.0%)	49 (77.8%)	44 (75.9%)	16 (48.5%)	11 (32.4%)	4 (14.3%)	1 (3.8%)	1 (10.0%)	1 (7.7%)	0 (0.0%)	0 (0.0%)

Extreme Disadvantage Scenario

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5	Layer 6	Layer 7	Layer 8	Layer 9	Layer 10	Layer 11	Layer 12	Layer 13	Layer 14
Number of applicants on the frontier	96	134	174	208	228	245	282	289	261	231	192	137	76	16
Number & percent minority	63 (65.6%)	69 (51.5%)	78 (44.8%)	66 (31.7%)	45 (19.7%)	28 (11.4%)	14 (5.0%)	5 (1.7%)	2 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Number & percent of minorities disadvantaged	63 (100%)	69 (100%)	78 (100%)	66 (100%)	45 (100%)	28 (100%)	14 (100%)	5 (100%)	2 (100%)	0 (NA)	0 (NA)	0 (NA)	0 (NA)	0 (NA)

Note: Table presents results from the data envelopment analysis peeling, separately for the no-discrimination scenario, as well as the moderate and extreme disadvantage scenarios. Each layer is the DEA frontier for the set of observations excluding all previous layers.

applicants in layers L through K as non-disadvantaged. The value L can be viewed as a parameter of the model: a value of L that is too small may under-classify applicants as disadvantaged and a value of L that is too large may over-classify applicants as disadvantaged. For our analysis, we set L to the minimum numbered layer for which the share of minority applicants within the layer is less than or equal to the share of minority applicants in the overall population. This value of L will classify borrowers as disadvantaged for each layer until the share of minority applicants in the layer is lower than the share of minority borrowers in the overall population.[24] Appendix C updates the algorithm from Richard S. Barr, Matthew L. Durchholz, and Lawrence Seiford (2000) for our analysis.

We look again at the example in Figure 2 for intuition. In the figure, there are 12 minority borrowers and 40 total borrowers, so the overall population share of minority borrowers is 30 percent. In the first layer 4 of 5 (80 percent) applicants are minorities, as indicated by filled versus unfilled circles, so we continue. In the second and third layers, 4 of 6 (66.67 percent) and 2 of 4 (50 percent) applicants are minorities, so we continue. When we reach the fourth layer, 1 of 5 (20 percent) applicants is a minority. This is less than the overall population share, so we stop. Minority applicants in layers 1-3 are classified as disadvantaged and the other minority applicants are classified as not disadvantaged.

Using this simple stopping criteria, we classify minority applicants in each of the three scenarios created by our DGP. Table 8 shows the resulting confusion matrices. The top left corners and bottom right corners of the matrices are the numbers of applicants for which the DEA peel

[24] One potential drawback of this simple heuristic is that the absence or presence of a single loan in the data could change the classification of an entire peel of borrowers. This potential stability problem is particularly pronounced when the share of DMUs in each layer is relatively high (e.g., because the dataset is small). An additional within-layer ranking technique (for an example, see Richard S. Barr, Matthew L. Durchholz, and Lawrence Seiford (2000) with a different cut-off metric could be used to increase the stability of the predictions.

procedure "correctly" identified disadvantage or the absence of disadvantage. The bottom left and top right corners of the matrices are the applicants for which the procedure "incorrectly" identified disadvantage or the absence of disadvantage. The tables show that the peels algorithm classifier correctly predicted disadvantage status for 68.4 percent, 81.9 percent, and 86.8 percent of applicants in each scenario, respectively.

Table 8: Example Confusion Matrix for Multi-peel Data Envelopment Analysis

No Discrimination Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	0 (0.0%)	117 (31.6%)
Predicted Not Disadvantaged	0 (0.0%)	253 (68.4%)

Moderate Disadvantage Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	190 (51.4%)	33 (8.9%)
Predicted Not Disadvantaged	34 (9.2%)	113 (30.5%)

Extreme Disadvantage Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	321 (86.8%)	0 (0.0%)
Predicted Not Disadvantaged	49 (13.2%)	0 (0.0%)

Note: Table presents classification results from the DEA-peels based classifier. A minority borrower is predicted to be disadvantaged if they fall on a layer before the first layer where the share of minority applicants within the layer is less than or equal to the share of minority applicants in the overall population.

ii) Intensity cutoff classifier

Our second method for identifying disadvantaged borrowers uses only the original pricing intensity scores. While disadvantaged borrowers may not fall on the frontier because of similarly-situated borrowers that received even more disadvantage, they should still lie closer to the frontier (i.e., have higher pricing intensity scores) than similarly-situated non-disadvantaged borrowers. Building on this simple intuition, the intensity cutoff method simply predicts borrowers above a certain intensity score threshold ("cutoff") to be disadvantaged and those below the cutoff to be non-disadvantaged.

As with the number of peels, this cutoff point can be viewed as a parameter of the model: a cutoff point that is too low may over-classify applicants as disadvantaged and a cutoff point that is too high may under-classify applicants as disadvantaged. For this analysis, we use the average intensity of non-minority borrowers. A minority borrower will be predicted to be disadvantaged, then, if and only if their pricing intensity is higher than the average non-minority borrower pricing intensity. As was true for the peeling approach, this cutoff would clearly classify a number of non-minority borrowers as disadvantaged if applied to them; however, as discussed above, a number of non-minority borrowers have high pricing intensities because there are no similarly-situated disadvantaged minority borrowers. Table 9 presents the resulting confusion matrices. The table shows that the intensity cutoff classifier correctly predicted disadvantage status for 54.3 percent, 82.7 percent, and 94.3 percent of applicants in each scenario, respectively.

iii) Cross-Intensity Method

Our third method is built on the DEA concept of cross-efficiency. DEA cross-efficiency was first introduced by Thomas R. Sexton et al. (1986) and greatly expanded on by John R. Doyle and Rodney H. Green (1994). Whereas efficiency (in our terminology, pricing intensity) scores for a given DMU are calculated using the optimal weights for *this* DMU, cross-efficiency (henceforth, cross-intensity for consistency) is calculated from the scores generated using *other* DMUs' weights. This method has been used and expanded in a number of contexts including preference voting (Rodney Green et al. 1996), portfolio selection in the Korean stock market (Sungmook Lim et al. 2014), and ranking countries' Olympic performances (Jie Wu et al. 2009).

Table 9: Example Confusion Matrix for DEA Intensity Cutoff

No Discrimination Scenario

	Disadvantaged	Not Disadvantaged
Predicted Disadvantaged	0 (0.0%)	169 (45.7%)
Predicted Not Disadvantaged	0 (0.0%)	201 (54.3%)

Moderate Disadvantage Scenario

	Disadvantaged	Not Disadvantaged
Predicted Disadvantaged	196 (53.0%)	36 (9.7%)
Predicted Not Disadvantaged	28 (7.6%)	110 (29.7%)

Extreme Disadvantage Scenario

	Disadvantaged	Not Disadvantaged
Predicted Disadvantaged	349 (94.3%)	0 (0.0%)
Predicted Not Disadvantaged	21 (5.7%)	0 (0.0%)

Note: Table presents classification results from the DEA intensity (efficiency) cutoff classifier. A minority borrower is predicted to be disadvantaged if and only if their pricing intensity is higher than the average non-minority borrower pricing intensity.

Returning to the notation and terminology used in our prior equations, we can find the pricing intensity for loan j using the weights for DMU i:

$$PricingIntensity_{ij} = \frac{\sum_{r=1}^{3} \gamma_{ri} y_{rj}}{\sum_{p=1}^{m} \delta_{pi} x_{pj} + +\delta_{0j}} \tag{4}$$

Where again y_{rj} is output r for pricing decision j, x_{pj} is input p for pricing decision j, γ_{ri} is the weight for output r for pricing decision i, and δ_{pi} is the weight for input p for pricing decision i. The traditional DEA intensity scores are simply $PricingIntensity_{jj}$, the diagonal elements of an

$n \, x \, n$ matrix. The cross-intensity for pricing decision j is calculated as the average of the intensity scores calculated using all other DMUs' weights, the matrix's column average.[25] So:

$$CrossIntensity_j = \frac{\sum_i PricingIntensity_{ij}}{n} \quad (5)$$

Here, n is the number of pricing decisions. However, the weights which maximize a particular pricing decision's intensity may not be unique, and the choice made from the set of non-unique optimal weights impacts the cross-intensity scores of the other pricing decisions (Sexton et al. 1986). A number of works attempt to resolve this issue by introducing a secondary goal. Most notably, Doyle and Green (1994) introduced the notion of "benevolent", "aggressive", and "arbitrary" formulations. In the aggressive (benevolent) formulation, the DMU's secondary goal is to make the cross-intensities of the other DMUs as small (large) as possible. The arbitrary formulation, by contrast, does not introduce an explicit secondary goal, generally leading to cross-intensity scores somewhere between the aggressive and benevolent formulations, depending on which of the solutions the linear programming algorithm finds. As neither the aggressive nor benevolent formulation are clearly preferable for the fair lending analyst, we use the arbitrary formulation. Further work could consider a secondary goal for this use case.

Cross-intensity scores have at least two theoretical advantages over traditional intensity scores for our analysis. First, because traditional intensity is calculated relative to a small number of peers, the DEA cutoff classifications based on traditional intensity may be unstable. For example, if a pricing decision references a single other decision on the frontier and that second

[25] As noted in Doyle and Green (1994), other measures than the mean (e.g., median, minimum, variances, etc.) could be used in certain contexts. Here we use the mean, which is widely used in the literature. Cross-intensity can be calculated including the traditional DEA intensity score for the DMU ("self-appraisal") or excluding this score. For our analysis, we choose to include these scores.

decision is removed, the intensity score of the first pricing decision will increase, perhaps dramatically. DEA cross-intensity, on the other hand, is calculated using the weights of every pricing decision in the data. Therefore, removing a single pricing decision will generally have a less sizeable impact and the classifications will be more stable. Second, as noted in Doyle and Green (1994), cross-intensity scores can be used to solve the “maverick” problem of DMUs receiving high intensity scores due to relatively unique weight sets. The fair lending analyst testing for pricing discrimination would be more interested in finding extreme observations in terms of outputs (i.e., high pricing compared to similar applicants), rather than simply inputs (i.e., very unusual applicants). Cross-intensity’s relative de-emphasis on applicants that have high pricing intensities simply because there are no similar applicants may, therefore, be preferable.

Table 10 shows the resulting cross-intensity in our scenarios. The table shows that in the no disadvantage scenario, minority applicants received slightly less cross-intense pricing than non-minority applicants (0.0042 less for minorities). Consistent with the DGP, in the moderate disadvantage scenario, minority applicants received more cross-intense pricing than non-minority applicants (0.0080 greater for minorities). The difference is larger in the extreme disadvantage scenario (0.0759 greater for minorities).

Table 10: Results Using Data Envelopment Analysis Cross-intensity

	No Disadvantage Scenario	**Moderate Disadvantage Scenario**	**Extreme Disadvantage Scenario**
Non-minority Applicants: Average Cross-Intensity	0.919	0.846	0.710
Minority Applicants: Average Cross-Intensity	0.915	0.854	0.786

Note: Table presents the average pricing cross-intensity for minorities and non-minorities, separately by discrimination scenario.

Table 11 again incorporates information the fair lending analyst would not have, displaying the average cross-intensity separately by the channel of disadvantage for the moderate and extreme disadvantage scenarios. In the moderate disadvantage scenario, the cross-intensity scores for disadvantaged applicants are generally higher than for others. The average cross-intensity for those disadvantaged in each of the three channels was between 0.866 and 0.886, whereas the average cross-intensity scores for non-disadvantaged minorities was only 0.823.

Table 11: Results for minorities using Data Envelopment Analysis cross-intensity

Moderate Disadvantage Scenario

Applicants	Average disadvantage for disadvantaged minorities	Average Cross-Intensity
76 Minorities disadvantaged through rate	45.33 bps	0.886
68 Minorities disadvantaged through points	1.92 pts	0.873
80 Minorities disadvantaged through fees	$4,572.39	0.866
146 Minorities not disadvantaged	NA	0.823

Extreme Disadvantage Scenario

Applicants	Average disadvantage for disadvantaged minorities	Average Cross-Intensity
120 Minorities disadvantaged through rate	121.48 bps	0.803
117 Minorities disadvantaged through points	4.78 pts	0.798
133 Minorities disadvantaged through fees	$11,433.74	0.759
0 Minorities not disadvantaged	NA	NA

Note: Table presents cross-intensities for minority borrowers, separately for the moderate and extreme disadvantage scenario. Each row includes borrowers who were disadvantaged through a particular channel.

We next use these cross-intensity scores to generate a binary classifier. Again, using a cutoff, we predict minority applicants with cross-intensity scores greater than the cutoff to be disadvantaged. This cutoff point is set to the average cross-intensity of non-minority borrowers. A minority borrower will be predicted to be disadvantaged, then, if and only if their cross-intensity is higher than the average non-minority borrower cross-intensity. Table 12 shows the resulting confusion matrices. The cross-intensity classifier correctly predicts disadvantage status for 60.3 percent, 80.8 percent, and 90.0 percent of applicants in each scenario, respectively.

Table 12: Example Confusion Matrix for Cross-intensity Data Envelopment Analysis

No Discrimination Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	0 (0.0%)	147 (39.7%)
Predicted Not Disadvantaged	0 (0.0%)	223 (60.3%)

Moderate Disadvantage Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	180 (48.6%)	27 (7.3%)
Predicted Not Disadvantaged	44 (11.9%)	119 (32.2%)

Extreme Disadvantage Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	333 (90.0%)	0 (0.0%)
Predicted Not Disadvantaged	37 (10.0%)	0 (0.0%)

Note: Table presents classification results from the DEA cross-intensity (cross-efficiency) cutoff classifier. A minority borrower is predicted to be disadvantaged if and only if their pricing cross-intensity is higher than the average non-minority borrower pricing cross-intensity.

iv) APR-based regression classifier

Finally, we attempt to classify borrowers using an extension of our APR-based regression analysis. We first construct the APR regression model discussed in the prior section. Then, we set the indicator on minority status to 0 for minority borrowers and predict APR according to the previously estimated model. This gives us a prediction of the APR that would be assigned to each minority applicant if they had the same input characteristics but were treated as non-minorities.

We then use the difference between the actual APR and the "treated as non-minority" predicted APR to determine whether the applicant was disadvantaged. For this analysis, borrowers with a positive difference (i.e., prediction < actual) are predicted to be disadvantaged.[26] Although it is not an entirely free parameter, this "difference cutoff point" could be adjusted to prevent over- or under-classifying borrowers as disadvantaged. Table 13 shows the resulting confusion matrices. The table shows that the APR-based regression classifier correctly predicted disadvantage status for 27.8 percent, 72.7 percent, and 100.0 percent of applicants in each scenario, respectively.

Table 13: Example Confusion Matrix for APR-based Regression

No Discrimination Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	0 (0.0%)	267 (72.2%)
Predicted Not Disadvantaged	0 (0.0%)	103 (27.8%)

Moderate Disadvantage Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	221 (59.7%)	98 (26.5%)
Predicted Not Disadvantaged	3 (0.8%)	48 (13.0%)

Extreme Disadvantage Scenario

	Disadvantaged	**Not Disadvantaged**
Predicted Disadvantaged	370 (100.00%)	0 (0.0%)
Predicted Not Disadvantaged	0 (0.0%)	0 (0.0%)

Note: Table presents classification results from the APR-regression based classifier. A minority borrower is predicted to be disadvantaged if and only if their individual error term from the regression described in Table 5 is smaller than the coefficient on minority status.

[26] In this simple linear model, then, a minority borrower is harmed if their individual error term is smaller than the coefficient on minority status.

v) Comparison using the Receiver Operating Characteristic curve

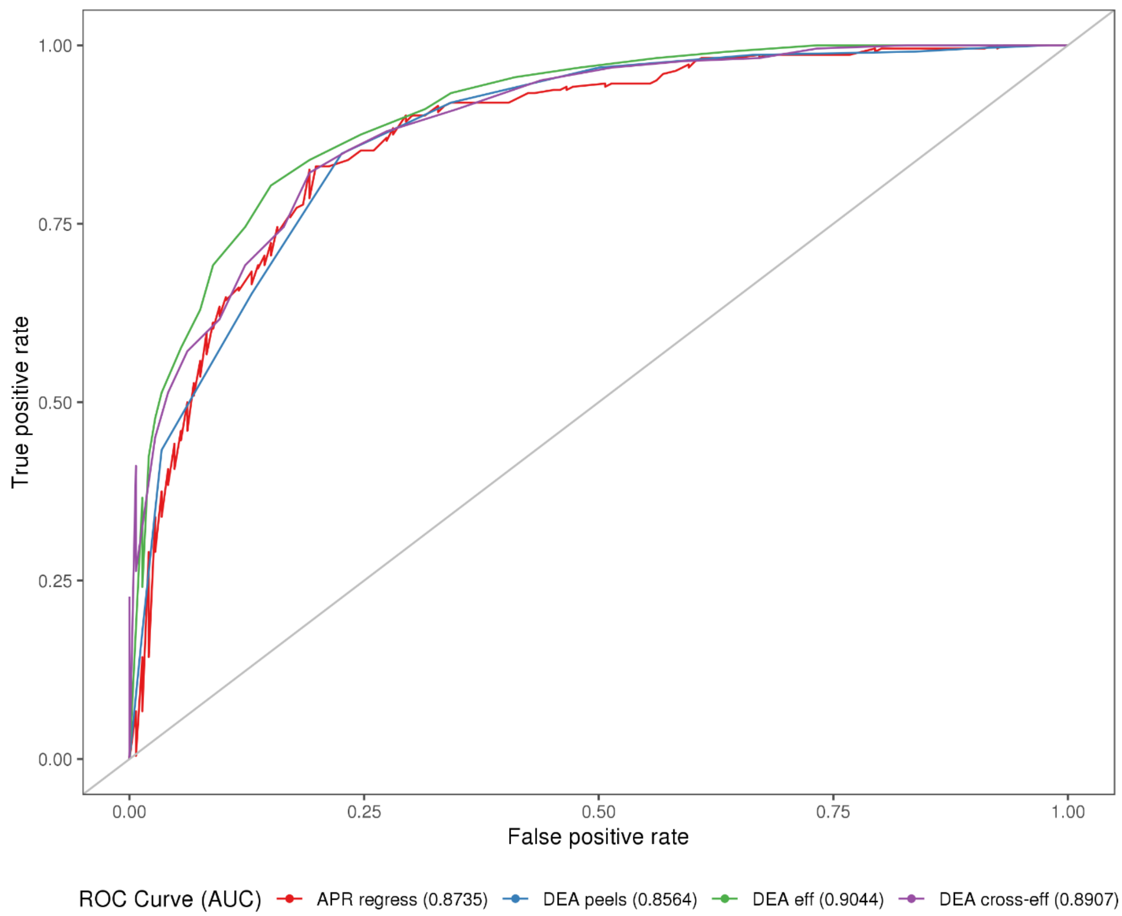
While the confusion matrices above provide one measure of accuracy for each prediction method according to the adjustable parameter for each model (number of peels, price intensity cutoff, cross-intensity cutoff, and predicted/actual difference cutoff), it is also important to assess each model's ability to classify applicants across different parameter values. One way of doing so is to plot the Receiver Operating Characteristic (ROC) curve. The ROC curve is constructed by adjusting the "parameter" for each model, then plotting the relationship between the false positive rate and the true positive rate over the range of possible parameter values. Figure 3 shows the ROC curves for the multi-peel DEA classifier, the intensity cutoff DEA classifier, the cross-intensity cutoff DEA classifier, and the APR-based regression classifier in the moderate disadvantage scenario.[27] The AUC is a measure of how well the model classifies applicants. An AUC value of 1 (or 100%) reflects perfect classification, and a value of 0.5 (or 50%) suggests that the classifier is only as good as a random guess. The figure shows the AUC for each of the four methods is similar, though the pricing intensity cutoff measures perform the best. The DEA pricing intensity cutoff method had the highest AUC (0.9044), followed by the cross-intensity cutoff (0.8907), then the APR-based regression method (0.8735), and the DEA peels method (0.8564).

V. Conclusion

Analyzing pricing decisions on mortgages for fair lending purposes is challenging given the tradeoffs of different components of price. In this study, we explored the merits of using DEA as an alternative to the standard single-equation APR approach. DEA is consistent with the historical approach in fair lending of focusing treatment on similarly-situated applicants.

[27] The moderate disadvantage scenario is the only scenario where some minorities are disadvantaged and others are not disadvantaged, a necessary condition for a useful AUC measure.

Figure 3: Receiver Operating Characteristic (ROC) curves for four methods of classification



Note: Figure shows Receiver Operating Characteristic curves, measuring the performance of four methods in identifying disadvantaged minority borrowers: (1) APR-regression based classifier, (2) a DEA-peels based classifier, (3) a DEA intensity (efficiency) cutoff classifier, and (4) a DEA cross-intensity (cross-efficiency) classifier.

Additionally, the technique does not make any distributional assumptions or assumptions about the tradeoffs between rate, points, and fees *a priori*, rather it identifies optimal combinations of rate, points, and fess empirically.

To assess a DEA approach, we used simulated data of a Retail lender using risk-based pricing, then generated two discrimination scenarios. The DEA identified patterns of discrimination in each of the two discrimination scenarios and did not incorrectly identify a pattern of discrimination in the original data. Finally, while the initial DEA frontier under-identified disadvantaged minorities, three DEA-based classifiers – a peeling-based classifier and two pricing intensity cutoff classifiers – had similar or higher AUC scores in comparison to an APR-based regression approach when attempting to identify disadvantaged minorities. These results are very promising for fair lending analysis. More broadly, our results suggest DEA may be useful in a wide variety of settings to detect discrimination in which tradeoffs exist between multiple outcome measures.

References

Avkiran, Necmi & Lin Cai (2014). Identifying distress among banks prior to a major crisis using non-oriented super-SBM. *Annals of Operations Research*, *217*(1), 31-53.

Banker, Rajiv D., Abraham Charnes, & William W. Cooper (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078-1092.

Banker, Rajiv D. & Richard C. Morey (1986). The use of categorical variables in data envelopment analysis. *Management science*, *32*(12), 1613-1627.

Barr, Richard S., Matthew L. Durchholz, & Lawrence Seiford (2000). Peeling the DEA onion: Layering and rank-ordering DMUs using tiered DEA. *Southern Methodist University Technical Report*, 5, 1-24. http://faculty.smu.edu/barr/pubs/layer.pdf.

Bartlett, Robert, Adair Morse, Richard Stanton, & Nancy Wallace (2019). Consumer-lending discrimination in the FinTech era (No. w25943). *National Bureau of Economic Research*. https://www.nber.org/papers/w25943.

Bhutta, Neil, Andrew C. Chang, Lisa J. Dettling, Joanne W. Hsu, & Julia Hewitt (2020). Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances. *FEDS Notes*, (2020-09), 28-2. https://www.federalreserve.gov/econres/notes/feds-notes/disparities-in-wealth-by-race-and-ethnicity-in-the-2019-survey-of-consumer-finances-20200928.htm

Bhutta, Neil & Aurel Hizmo (2021). Do Minorities Pay More for Mortgages? *Review of Financial Studies*, *34*(2), 763-789.

Black, Harold A., Thomas P. Boehm, & Ramon P. DeGennaro (2003). Is there discrimination in mortgage pricing? The case of overages. *Journal of Banking & Finance*, *27*(6), 1139-1165.

Boehm, Thomas P., Paul D. Thistle, & Alan Schlottmann (2006). Rates and race: An analysis of racial disparities in mortgage rates. *Housing Policy Debate*, *17*(1), 109-149.

Bougnol, Marie-Laure & José H. Dulá (2006). Validating DEA as a ranking tool: An application of DEA to assess performance in higher education. *Annals of Operations Research, 145*(1), 339-365.

Brown, Rayna (2006). Mismanagement or mismeasurement? Pitfalls and protocols for DEA studies in the financial services sector. *European Journal of Operational Research, 174*(2), 1100-1116.

Charnes, Abraham, William W. Cooper, & Edwardo Rhodes (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444.

Cheng, Ping, Zhenguo Lin, & Yingchun Liu (2015). Racial Discrepancy in Mortgage Interest Rates. *Journal of Real Estate Finance and Economics*, *51*(1), 101-120.

Courchane, Marsha & David Nickerson (1997). Discrimination resulting from overage practices. *Journal of Financial Services Research, 11*, 133-151.

Courchane, Marsha (2007). The pricing of home mortgage loans to minority borrowers: How much of the APR differential can we explain? *Journal of Real Estate Research*, *29*(4), 399-440.

Crawford, Gordon W. & Eric Rosenblatt (1999). Differences in the cost of mortgage credit implications for discrimination. *The Journal of Real Estate Finance and Economics*, *19*(2), 147-159.

Dietrich, Jason (2005). Under-specified Models and Detection of Discrimination: A Case Study of Mortgage Lending. *Journal of Real Estate Finance and Economics*, *31*(1), 83-105.

Dodd-Ramirez, Daniel, & Patrice A. Ficklin (2016). Redlining: CFPB and DOJ action requires BancorpSouth Bank to pay millions to harmed consumers. *Consumer Financial Protection Bureau.* https://www.consumerfinance.gov/about-us/blog/redlining-cfpb-and-doj-action-requires-bancorpsouth-bank-pay-millions-harmed-consumers/

Doyle, John R. & Rodney H. Green (1994). Efficiency and cross-efficiency in DEA: Derivations, meanings and uses. *Journal of the Operational Research Society*, *45*(5), 567-578.

Dyson, R. G., R. Allen, A. S. Camanho, V. V. Podinovski, C. S. Sarrico, & E. A. Shale (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research, 132*(2), 245-259.

Ebersberger, Bernd, Uwe Cantner, & Horst Hanusch (1997). Consumer choice and data envelopment analysis: the case of discrimination on housing markets. Paper presented at the *La methode DEA et l'analyse des performances et des organisations* symposium.

El Anshasy, Amany, Gregory Elliehausen, & Yoshiaki Shimazaki (2004). Mortgage brokers and the subprime mortgage market. *George Washington University Working Paper*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.8277&rep=rep1&type=pdf

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General), 120*(3), 253-281.

Ghent, Andra C., Rubén Hernández-Murillo, & Michael T. Owyang (2014). Differences in subprime loan pricing across races and neighborhoods. *Regional Science and Urban Economics*, *48*, 199-215.

Green, Rodney H., John R. Doyle, & Wade D. Cook (1996). Preference voting and project ranking using DEA and cross-evaluation. *European Journal of Operational Research*, *90*(3), 461-472.

Hanson, Andrew, Zackary Hawley, Hal Martin, & Bo Liu (2016). Discrimination in mortgage lending: Evidence from a correspondence experiment. *Journal of Urban Economics*, *92*, 48-65.

Holmes, Andrew & Paul Horvitz (1994). Mortgage Redlining: Race, risk, and demand. *Journal of Finance*, *49*(1), 81-99.

Kearns, Deborah & Barbara Marquand (2019). Mortgage Closing Costs: What They Are and How Much You'll Pay. *NerdWallet.* https://www.nerdwallet.com/article/mortgages/closing-costs-mortgage-fees-explained.

Kleiner, Morris M. & Richard M. Todd (2009). Mortgage broker regulations that matter: Analyzing earnings, employment, and outcomes for consumers. In *Studies of Labor Market Intermediation* (pp. 183-231). University of Chicago Press.

Ladd, Helen F. (1998). Evidence on Discrimination in Mortgage Lending. *Journal of Economic Perspectives, 12*(2), 41-62.

Lake, Rebecca (2019). When Is Paying a Loan Origination Fee Worth It? *U.S. News & World Report.* https://loans.usnews.com/articles/when-is-paying-a-loan-origination-fee-worth-it.

Lewis, Herbert F. & Thomas R. Sexton (2004). Data envelopment analysis with reverse inputs and outputs. *Journal of Productivity Analysis*, *21*(2), 113-132.

Lim, Sungmook, Kwang Wuk Oh, & Joe Zhu (2014). Use of DEA cross-efficiency evaluation in portfolio selection: An application to Korean stock market. *European Journal of Operational Research*, *236*(1), 361-368.

Lober, Gerrit & Matthias Staat (2010). Integrating categorical variables in Data Envelopment Analysis models: A simple solution technique. *European Journal of Operational Research*, 202(3), 810-818.

Longhofer, Stanley D. (1998). Measuring Pricing Bias in Mortgages. *Economic Commentary*, Federal Reserve Bank of Cleveland, August 1.

Longhofer, Stanley D. & Paul S. Calem (1999). Mortgage brokers and fair lending. *Economic Commentary*, Federal Reserve Bank of Cleveland, May 15.

Longhofer, Stanley D. (2000). APR flawed for fair-lending analysis. *Regulation*, 23, 12.

Munnell, Alicia H., Geoffrey M.B. Tootell, Lynn E. Browne, & James McEneaney (1996). Mortgage Lending in Boston: Interpreting HMDA Data. *American Economic Review*, *86*(3), 25-53.

Nothaft, Frank E. & Vanessa G. Perry (2002). Do mortgage rates vary by neighborhood? Implications for loan pricing and redlining. *Journal of Housing Economics*, 11(3), 244-265.

Pastor, Jesus T. & Jose L. Ruiz (2007). Variables with negative values in DEA. In *Modeling data irregularities and structural complexities in data envelopment analysis* (pp. 63-84). Springer, Boston, MA.

Ross, Stephen L. & John Yinger (2002). *The color of credit: Mortgage discrimination, research methodology, and fair-lending enforcement*. MIT press.

Sarkis, Joesph (2007). Preparing your data for DEA. In *Modeling data irregularities and structural complexities in data envelopment analysis* (pp. 305-320). Springer, Boston, MA.

Sexton, Thomas R., Richard H. Silkman, & Andrew J. Hogan (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, *1986*(32), 73-105.

Tootell, Geoffrey M.B. (1996). Redlining in Boston: Do Mortgage Lenders Discriminate Against Neighborhoods? *Quarterly Journal of Economics*, *111*(4), 1049-1079.

Wen, Meilin (2015). *Uncertain data envelopment analysis*. Springer.

Wu, Jie, Liang Liang, & Yao Chen (2009). DEA game cross-efficiency approach to Olympic rankings. *Omega*, *37*(4), 909-918.

Yezer, Anthony M. (2010). A review of statistical problems in the measurement of mortgage market discrimination and credit risk. *Available at SSRN 1684216*. https://ssrn.com/abstract=1684216

Zhang, David H. & Paul Willen (2020). Do Lenders Still Discriminate? A Robust Approach for Assessing Differences in Menus. *Available at SSRN 3688604.* https://ssrn.com/abstract=3688604

Appendix A: Data Generating Process

A rate sheet is typically the starting point for any risk-based pricing, and the foundation of the rate sheet is a rate/point schedule. Rate/point schedules present the number of discount points the lender is willing to accept for a variety of interest rates. Applicants with a longer time horizon may wish to pay discount points to buy down the rate, while applicants with short-term liquidity constraints may be willing to pay a higher rate in return for rebate points. We assume the lender is indifferent between all rate/point combinations on a given rate/point schedule.

Rate sheets also include risk-related adjustments to rate and/or discount points called loan-level price adjusters or LLPAs. The values of these LLPAs are applicant specific, and capture the risk profile of the applicant. Typically, lenders apply these adjustments to the base rate/point schedule, thereby altering the number of discount points an applicant must pay for a given rate (point adjuster) or the rate an applicant can get for a given number of discount points (rate adjuster). Alternatively, some lenders treat these risk adjusters as origination costs the applicant must pay at closing. In addition to the rate and discount points, it is also common for lenders to charge origination points to cover processing costs, and fees to cover third-party vendor costs.

When an applicant applies for a mortgage, there is typically an initial discussion with the lender about her credit needs, and what type of product might meet those needs. Based on the information gathered during this discussion, the lender identifies the specific rate sheet appropriate for the transaction. The specific rate sheet applied will be based on a number of product and applicant characteristics such as loan type, product, program, amortization type, term, lock-in period, date, and geography. Once the rate sheet is identified the lender then applies the applicant's specific LLPAs to generate the rate/point schedule relevant to the specific applicant. The lender and applicant then negotiate the specific rate/points combination for the loan.

Rate and discount points are determined simultaneously, so we represent this process as a system of two equations. The first equation (A1a) is the rate equation where rate depends on the rate/point schedule from the rate sheet along with any rate sheet adjustments to rate. The indicator variable $I_{dpts,k}$ identifies which of the k = 1, .., K specific rate/point pairs the applicant chose from the relevant rate/point schedule. $\alpha_{a,k}$ conveys the specific amount of discount points corresponding to the rate/point pair chosen. $\alpha_{a,k}$ is indexed by the commitment period a = 15, 30, 45 days to show that there are different rate/point schedules by commitment period[28]. $I_{dpts,k}$ and $\alpha_{a,k}$ therefore reflect the applicant's decisions on rate lock period and whether to buy up or buy down the rate. RA_g is an indicator variable denoting whether rate adjuster g = 1,…G is applicable to the specific applicant and β_g is the specific value of the rate adjuster. Although pricing is separately determined for each applicant, we exclude the *i* subscript throughout for ease of exposition.

The second equation (A1b) is the discount points equation. Discount points paid depend on the rate/point schedule from the rate sheet plus any rate sheet adjustments to points. Similar to equation 1a, the indicator variable, $I_{R,k}$ identifies which of the k = 1, .., K specific rate/point pairs the applicant chose from the relevant rate/point schedule; $\gamma_{a,k}$ conveys the specific rate corresponding to the rate/point combination chosen; PA_g is an indicator variable denoting whether point adjuster h = 1,…H is applicable to the specific applicant; and δ_g is the specific value of the point adjuster.

[28] Although there may be separate rate/point schedules for different loan types, products, programs, amortization types, terms, dates, and geographies, we only allow for different rate/point schedules for rate lock period for ease of exposition.

It should be noted that all point LLPAs are implicitly included in $\alpha_{a,k}$ in equation A1a and all rate LLPAs are implicitly included in $\gamma_{a,k}$ in equation A1b. Therefore, equations A1a and A1b are the inverse of each other and capture the same post-LLPA, rate/point relationship. We include both equations to allow for the possibility that lenders can apply discretionary adjustments to rate, points, or both.

$$R = \sum_{a=15}^{45}\left\{\sum_{k=1}^{K}(\lambda_{a,k}\alpha_{a,k} + \sum_{g=1}^{G}\beta_g * RA_g) * I_{dpts,k}\right\} * I_{lock,a} \quad \text{(A1a)}$$

$$DPTS = \sum_{a=15}^{45}\left\{\sum_{k=1}^{K}(\theta_{a,k}\gamma_{a,k} + \sum_{h=1}^{H}\delta_h * PA_h) * I_{R,k}\right\} * I_{lock,a} \quad \text{(A1b)}$$

where,

R = Interest Rate

$I_{dpts,k}$ = 0/1 Indicator variable for each discount point value k on the rate sheet

$\alpha_{a,k}$ = represents each discount point value k on the rate sheet

$\lambda_{a,k}$ = transformation of points into rate (percent)

$I_{lock,a}$ = 0/1 Indicator variable for each lock-in period a on the rate sheet

RA_g = 0/1 Indicator variable for each rate adjustment g on the rate sheet

β_g = represents each rate adjustment value g from the rate sheet

DPTS = Discount Points

$I_{R,k}$ = 0/1 Indicator variable for each interest rate value k on the rate sheet

$\gamma_{a,k}$ = represents each interest rate value k on the rate sheet

$\theta_{a,k}$ = transformation of rate (percent) into points

PA_h = 0/1 Indicator variable for each point adjustment h on the rate sheet

δ_h = represents each point adjustment value h from the rate sheet

In addition to rate and discount points, lenders also charge origination points and third-party fees. Lenders often charge 1 origination point to cover administrative costs, and we take this as our policy as reflected by Equation (A2). Third party fees cover costs such as title search, credit report, property taxes, and flood insurance. Some of these fees will impact APR and others will not. We separate these fees out, because APR is one measure of price commonly analyzed in the fair lending literature. Throughout this study, we assume that lenders set non-APR-affecting, third party fees to just cover the costs of third-party vendors, and therefore do not allow for discretionary adjustments of these fees that increase lender's profits. Equation (A4) represents total closing costs, which are equal to fees that affect APR, fees that do not affect APR, discount points, origination points, and any down-payment from the applicant, net of any seller or lender contributions.[29][30] The final equation in our model (A5) is just an identity showing that for a given interest rate, the lender charges the applicant its reservation price, which is equal to discount points plus origination points.

$$OPTS = 1 \quad \text{(A2)}$$

$$CC = VC_{APR} + VC_{NonAPR} + DPTS + OPTS + DP - Z \quad \text{(A3)}$$

$$Price = DPTS + OPTS \quad \text{(A4)}$$

where,

CC	= Total Closing Costs
VC_{APR}	= Third-party vendor costs that affect APR
VC_{NonAPR}	= Third-party vendor costs that do not affect APR
DP	= Down-payment
Z	= Total seller and lender contributions to closing costs
Price	= Reservation price to purchase a loan with rate R

Equations A1-A4 reflect the lender's pricing policies and decision-making process with no discretion. Typically, however, lenders allow their loan officers discretion to deviate from the rate

[29] In reality, down-payment and loan amount are both endogenous and dependent on the applicants' wealth, liquid assets and preferences. However, we make the simplifying assumption here that both are exogenous. Loan amount is often either a rate or points adjuster on rate sheets and therefore may affect rate and points directly. Loan amount also typically affects rates and points indirectly through CLTV.

[30] These specifications do not include the likely tradeoffs in loan amount and possibly rate and discount points if the seller or lender contribute to closing costs.

sheet to address applicants' preferences on the components of price. As one example, an applicant may express a preference for paying no discount points. We incorporate that discretion here. We continue to impose the restriction that the lender just receives the reservation price for each loan, but now allow the source of price to vary. Specifically, discount points paid no longer have to equal the discount points from the rate sheet for a given rate and origination points paid no longer have to equal 1.[31] This adjustment incorporates the trade-offs between discount points and origination points into our model. Equations A1a, A1b, A5, and A6 present our new system. Even though we have introduced some discretion into the model, the lender still receives only the reservation price as indicated by equation (A6).

$$DPTS_{paid} = Price_{a,k} - OPTS_{paid} \quad \text{(A5)}$$

$$CC = NVC_{APR} + NVC_{NonAPR} + DPTS_{paid} + OPTS_{paid} + DP - Z \quad \text{(A6)}$$

[31] Although possible, we do not allow for lenders to adjust third party fees to make up for a shortfall in a different component of price.

Appendix B: Rate Sheet and LLPAs for 30-Year Fixed Rate Products

	Lock Days		
	15 Day	30 Day	45 Day
4.250%	2.350	3.000	3.450
4.375%	1.850	2.500	2.950
4.500%	1.350	2.000	2.450
4.625%	0.850	1.500	1.950
4.750%	0.350	1.000	1.450
4.875%	(0.150)	0.500	0.950
5.000%	(0.650)	0.000	0.450
5.125%	(1.150)	(0.500)	(0.050)
5.250%	(1.650)	(0.875)	(0.550)
5.375%	(2.150)	(1.125)	(1.050)
5.500%	(2.650)	(1.375)	(1.550)

Adjustments:

Fico>=700 and CLTV>80	1.25 to points
Fico<700 and CLTV<=80	0.75 to points
Fico<700 and CLTV>=80	2.00 to points
Loan amount <= 200,000	1.5 to points

Appendix C: Peeling Algorithm[32]

(1) Initialize: $t \leftarrow 1,\ D^{[1]} \leftarrow D$

(2) Apply a DEA model to the DMUs in set $D^{[1]}$ to identify ${E^*}_{[1]}$

(3) While $M({E^*}_{[t]}) > M(D)$ do:

a. ${I^*}_{[t]} \leftarrow \mathrm{D[t]} - {E^*}_{[t]}$

b. $t \leftarrow t + 1$

c. $D^{[t]} \leftarrow {I^*}_{[t-1]}$

d. Apply a DEA model to the DMUs in set $D^{[t]}$ to identify ${E^*}_{[t]}$

Where t is a tier index, ${E^*}_{[t]}$ and ${I^*}_{[t]}$ are the sets of efficient and inefficient DMUs (pricing decisions) on tier t that makeup the dataset $D^{[t]}$. $M({E^*}_{[t]})$ and $M(D)$ are the share of pricing decisions on frontier t and the share of pricing decisions in the entire dataset that are for minority borrowers.

[32] This algorithm simply updates (for use in fair lending analyses) the algorithm introduced in Richard S. Barr, Matthew L. Durchholz, and Lawrence Seiford (2000).