Gender Identity and Access to Higher Education

I. Maupin*
West Virginia University

Bryan C. McCannon†
West Virginia University

August 30, 2023

Abstract

Our research examines the potential gender identity discrimination within higher education. An audit study was conducted by sending emails to admission counselors, where the messages varied in the inclusion of gender pronouns in the signature line. The results indicate a higher response rate for emails which included preferred pronouns, with a response rate increase of approximately four percentage points, regardless of the type of pronoun used. This suggests a preference for students that are more progressive in their thinking. We engage in text analysis and show that responses to inquiries with pronouns received more friendly responses receiving more use of exclamation marks, emojis/emoticons, and from a topic modeling algorithm were less likely to be strictly replies explaining the admission process. Finally, we apply machine learning to identify key institution attributes that are useful in predicting heterogeneous responses, and to identify the attributes of institutions where negative discrimination is likely to occur.

JEL Codes: J16, I24, K38
Keywords: admissions; audit study; discrimination; gender; higher education; machine learning; neo-pronouns; pronouns; Random Forest

1 Introduction

The recognition of the non-binary nature of gender identity has increased and there is concern over discrimination experienced by non-binary individuals [Waite, 2021]. One critical area where discrimination may occur is in higher education, particularly in the admissions process [Lawrence and Mckendry, 2019]. In light of heightened legal scrutiny regarding discrimination in admissions, this study specifically explores the potential impact of gender identity on the admissions process.

For individuals whose gender identity does not align with the traditional binary labels of "male" or "female", the prevalence of gendered language in English requires the use of neo-pronouns for clear self-expression. Non-binary individuals can specify their preferred gender pronouns in their communications, and those who identify with the binary labels can also choose to provide their preferred pronouns as a sign of inclusivity and alignment. By including preferred pronouns, whether they are binary or non-binary, individuals can communicate their gender identity to others.

*Department of Economics, John Chambers College of Business and Economics, 83 Beechurst Avenue, Morgantown WV 26506, im00010@mix.wvu.edu

†Corresponding author: Associate Professor, Department of Economics, John Chambers College of Business and Economics, 83 Beechurst Avenue, Morgantown WV 26506, bryan.c.mccannon@gmail.com We appreciate valuable discussions with Nathaniel Burke, Alicia Plemmons, Josh Martin, Zach Rodriguez, Cameron Tilley, and seminar participants at the 2023 ScaledUp workshop and Association of Private Enterprise Education annual meetings.

During the admissions process, potential applicants typically communicate with admission counselors employed by the higher education institution. These counselors are tasked with helping in the application process and promoting the institution. Hence, admission counselors affect the transaction costs associated with applying to college. Since communication of preferred pronouns will reach these admission counselors, there is a risk of discrimination based on gender identity. Consequently, we ask whether gender identity affects the admissions process.

If an institution of higher education, represented by the admission counselors as its agents, values progressively-minded individuals, then the communication of pronouns can be viewed favorably. If it specifically values non-binary students, then communication of a non-binary gender could also lead to enhanced support. On the other hand, if an institution's agent experiences a discriminatory disutility to either progressively-minded and/or non-binary individuals on campus, then they could choose not to lower the transaction costs by neglecting those potential applicants.

To investigate this question, we engage in an audit study. In it, we send emails to admission counselors at 500 randomly-selected higher education institutions in the United States. The message asked a simple question related to completing the application. The outcome variable of interest is whether the admission counselor responded to the request for help. In a within-subject design, treatments differ in the inclusion of preferred gender pronouns. The control does not include pronouns. Two treatments include (he/him) and (she/her) in the signature line. A final treatment includes neo-pronouns.

We find that there is, in fact, a *greater* response rate to messages which include preferred pronouns. Further, there is not a difference between treatments where traditional, binary pronouns are used and where neo-pronouns are included. We interpret these results as suggesting that agents of higher education institutions hold a preference for progressively-minded individuals. In addition, text analysis of responses to our message suggest more positive, friendlier replies when preferred pronouns are included. This includes heightened use of exclamation marks and emojis/emoticons. Topic modeling reveals a shift in content towards friendlier messages away from strictly factual replies regarding the admission process.

Further, higher education is a highly differentiated market. Institutions contrast on numerous attributes. We engage in machine learning to identify which attributes are most useful in predicting replies. We use this to guide our identification of heterogeneous treatment effects. We find that institutions not within the lowest quintile of undergraduate student population size, low retention rates, large proportions receiving Pell grants, and those located within a city are the type of institutions that exhibit the strongest preference for progressively-minded students. Relatedly, we use machine learning to hunt for a subsample of institutions who exhibit responses consistent with having a negative discriminatory taste against non-binary individuals. We find that those with low student-faculty ratios, a small proportion of students receiving financial aid, and being in the northeast part of the United States are those who most likely to discriminate against non-binary individuals.

Research on potential discrimination based on non-binary gender is sparse. To the best of our knowledge, ours is the first to explore its influence in higher education. Button et al. [2020] conducts an audit study which includes direct statements of transgender or non-binary (TNB) gender status in appointment requests to mental health facilities. They find reduced positive responses to requests made by TNB individuals. Granberg et al. [2020] conducts an audit study completing job applications. The treatment includes cover letters where the applicant discloses that they legally changed their first name, which is meant to signal transgender status. Transgender applications received a lower response rate.

Our results suggest that ideology may be more relevant than non-binary gender identity in higher education. Ideological bias in higher education has received attention. For one prominent example, Bar and Zussman [2012] provide evidence that faculty member's political affiliation influences grades awarded, especially for racial minority students. The reduction in ideological diversity among faculty on college campuses has been documented [Duarte et al., 2015] and the public's perceptions of research findings are influenced by their beliefs about the ideological bias of the researchers [MacCoun and Paletz, 2009]. The focus of the past research, though, has been on the faculty's ideology. To the best of our knowledge, our result is the first to document a preference for a specific ideology of students in the admission process.

Experiments on college admissions are also rare. A notable exception is Stewart and Uggen [2020]. They conduct an audit study completing applications. Treatments vary by the inclusion of criminal records. They find a greater rejection rate for applicants with a criminal history. We do not go as far as to fill out fictitious applications, but focus instead on pre-application communication where gender identity can be easily communicated. The closest paper to ours is Hanson [2017] where an audit study also sending emails to college admissions counselors is conducted. He varies the names differentiating distincively-Black from non-distinctively-Black names. Also, messages vary by gender and student quality (SAT scores and high school GPA).[1] Hanson [2017] finds that counselors reply at a slightly higher rate to students with traditional-female names, but display no differentiation on average between distinctively-Black and others regardless of student quality. Hanson [2017] does not explore non-binary gender status as we do.

Finally, research has explored the decision to identify with minority groups. Antman and Duncan [2015] evaluate survey data of multi-racial individuals separately reporting ancestry and self-identified race. They show that willingness to identify with an under-represented race diminishes when states ban affirmative action in public university admissions. Additionally, identifying with an over-represented race expands. The mitigation of gendered-language on exams has been shown to improve women's performance on quantitative questions [Cohen et al., 2023]. Aneja et al. [2023] explores demand for Black-owned businesses documenting that when owner's race becomes salient[2] demand increases and that this comes disproportionately from White consumers. While these efforts study the decision to identify with a particular race, we complement them by exploring gender identity.

[1] Data on popular names of baby-girls and baby-boys are consulted to select the names.
[2] Specifically, restaurants could choose to be designated as a Black-owned establishment on Yelp.

2 Theory

We present a simple theoretical framework to guide the analysis. Consider a representative university representative who receives an email inquiry from a prospective student. They may either respond to the message or ignore it. Choosing to respond to the message increases the chance that the prospective student applies and becomes a student on campus. Normalizing the utility to not responding equal to zero, the representative receives an expected payoff from responding equal to

$$u = \pi + \lambda + \nu - \kappa. \tag{1}$$

In this expected payoff function, their utility is comprised of four components.

First, $\kappa > 0$ is the cost associated with taking the time to respond to the message and help the prospective student in their application. Thus, the representative needs to be motivated to exert effort.

Second, they receive a benefit from increasing the applicant pool, $\pi > 0$. As an institution has more student demand, one can imagine there would be a corresponding increase in expected enrollment and, consequently, revenue. Alternatively, an increase in the applicant pool can lead to lower discount rates on tuition (again increasing expected university revenue) or allow the institution to be more selective in who it admits, improving its quality and reputation. Regardless, the admission counselor would care about expected revenue, quality, and reputation through job security and financial compensation.

Third, we hypothesize that the institution's representative cares about the ideology of the student body. The term λ is the additional value placed on having a liberally-minded, progressive individual in the applicant pool. Thus, one can think of π as the baseline value of a "neutral" applicant in the pool and λ as the additional benefit from that applicant being liberal. A representative who does not care about political/economic/social ideology has a value of $\lambda = 0$. One who values progressively-minded individuals has $\lambda > 0$, while a socially-conservative representative has $\lambda < 0$.[3]

Fourth, we also allow for the institution's representative to care about the gender identity of the applicant. The term ν is the additional expected benefit to having another non-binary individual on campus. We think of it as an expected benefit as the representative will be unsure of the gender identity and values of the prospective student. The representative may view there to be a positive benefit to the campus community/culture by having a diverse student body, for example. In this case $\nu > 0$. Alternatively, a "distaste" for non-binary individuals would have $\nu < 0$. By caring about the campus culture directly, the gender identity of the prospective student may matter to the representative.

[3]Our experiment does not provide a way to signal a conservative ideology. Therefore, we omit this from the model, but one could think of a negative value to λ being a conservatively-minded agent. In addition, the theoretical model does not differentiate between traditional male and female pronouns. One could reasonably expect that the updating of the agent's prior beliefs can differ in magnitude when a male applicant adds preferred pronouns than when a female applicant does. Since, as will be shown in the upcoming results, there is not a marked difference in response rates, we choose not to complicate the model by adding this distinction. Further, the simplistic model does not include the choice to include the message (taking it as exogenous). A fully-developed, cheap-talk signaling model would formalize the expectations of the applicant's ideology given the existence of the message being sent. Our reduced model can be thought of as considering the expected payoffs generated from the intervention.

In our experiment, which we describe in detail in the upcoming section, our control will be stripped of information that would allow the receiver to make assessments on ideology or gender identity. Thus, the expected utility of replying is $u(control) = \pi - \kappa$ and they respond only if $\pi \geq \kappa$.

In two treatments we add traditional gender pronouns to the message's signature line. In one treatment we add male pronouns and in the other we add female pronouns. We argue that this potentially provides information to the institution's representative on the prospective student's ideology. A cis-gender individual may add these traditional pronouns as a way to signal allegiance, support inclusivity, or virtue signal. Regardless, we argue that the inclusion of pronouns conveys information correlated with progressive-liberal ideology. Thus, the representative will respond to the message when $u(traditional) \geq 0$, or rather, when $\pi + \lambda \geq \kappa$. Consequently, the response rate difference between the control and these treatments assesses whether $\lambda \lesseqgtr 0$.

Finally, our last treatment includes neo-pronouns. We argue that this provides information to the institution's representative on both the ideology and the gender identity of the prospective student. Hence, the expected utility to responding to this message is $u(neo) = \pi + \lambda + \nu - \kappa$. Comparing the response rate in this treatment to the control assesses whether $\lambda + \nu \lesseqgtr 0$, while comparing its response rate to the traditional pronoun treatments asks whether $\nu \lesseqgtr 0$.

Hypothesis 1: *If higher education institutions prefer progressively-minded students, then the response rate to messages with traditional pronouns included will be higher than messages without pronouns included.*

Hypothesis 2: *If higher education institutions prefer diversity in gender identity, then the response rate to messages with neo-pronouns included will be higher than messages with traditional pronouns included.*

Hypothesis $2'$: *If higher education institutions have discriminatory preferences against non-binary individuals, then the response rate to messages with neo-pronouns included will be lower than messages with traditional pronouns.*

3 Methods

3.1 Sample

To select our sample of higher education institutions, we start with a list of all institutions in the United States who report data to the National Center for Education Statistics (NCES), which is part of the U.S.

Department of Education. We trim the list eliminating (1) all 2-year institutions, (2) specialized colleges[4], (3) institutions with less than 500 students, (4) military academies, and (5) closed institutions that no longer admit students. From this list, we select randomly a sample of 500 higher education institutions.

For each of these 500 schools, we visit its web site and collect email addresses for four admission counselors. The typical admissions office has numerous individuals employed in a variety of positions. We focus on admissions counselors as their job is to facilitate and recruit new students. They are usually the primary point of contact for applicants and, therefore, their willingness to help out an individual in their application can reduce the transaction cost to applying. We argue that admission counselors unwilling to help act as a barrier and can harm someone's pursuit of a higher education.

3.2 Audit Study

We sent an email to each counselor using dummy Gmail accounts. Each email sent included an identical subject line, "question". The message asks for help in the application. Specifically, the email reads,

> I am completing your admissions application, but I have a question. My family is currently in the process of moving, but it will be in a few months. Should I use my current address or the new address for filling out the application?
>
> Thank you.

The message ends with a signature line. Treatments differ only in the signature line, which includes the sender's name and preferred pronouns (or no pronouns for the control).

We choose four names to be the originators of the messages: Alex Johnson, Morgan Johnson, Sam Johnson, Taylor Johnson. We use the last name Johnson as it is a common last name. The four first names were selected from a list of the most common unisex names. Our goal was to use neutral names that did not strongly imply a gender identity so that only the pronoun usage conveyed the information we manipulated. Also, we do not think these are distinctively-Black (or Hispanic, Asian, etc.) names. Our focus is on gender identity rather than race-gender interactions.

We set up four dummy email accounts for our four fictitious individuals. The domain for each is gmail.com and the usernames are a.johnson32005, m.johnson42005, s.johnson42005, and t.johnson42005. We use the first initial to further avoid gender presumptions. The numbers added to the username are intended to suggest a high school student born in spring 2005 who would be a typical applicant.[5]

Our experiment is a within-subject design with four treatments for each institution.[6] The *Control* leaves the signature empty. The name is included but no pronouns are used. The *Male* treatment inserts (he/him)

[4] Our supplemental appendix lists fully the categories excluded, but common examples include architectural, culinary, medical, and theological institutions.

[5] To our surprise a.johnson42005 was already taken when we set up these accounts.

[6] Except for those universities which have limited contact information, which force us to have a slightly unbalanced panel. Also, it is within-subject if one takes of the institution as the subject (rather than the admission counselor).

in the signature. The *Female* treatment inserts (she/her) in the signature. Finally, the *Neo-Pronouns* treatment inserts the neo-pronouns (xe/xem) in the signature.

There are numerous neo-pronouns used in practice. While common, we chose not to use (they/them). Our reasoning was that cis-gendered individuals occasionally use these pronouns to support inclusivity and mitigate gendered language.[7] Consequently, we could not ensure that the communication of (they/them) would always be interpreted as a non-binary applicant. Someone may, for example, identify as female but prefer to reduce the use of gender in the English language by supporting the use of the singular 'they'. To mitigate this ambiguity, we use neo-pronouns which we expect more strongly convey a non-binary individual. The ones we use replace the gendered part of the pronoun with the letter X.

3.3 Protocol

We pre-registered our design with the American Economic Association and received Exempt certification from West Virginia University's Institutional Review Board before proceeding.[8] Our sample size selected was approximately twice as large as needed from a power calculation using the measured effect size from a 30-school pilot study we conducted prior to our experiment.

The emails were sent in 12 waves in November and December 2022. These months represent the typical time window for many applications as, frequently, early admission decisions come in October and regular admissions arrive in January (or later) of the following year. The four emails sent to each institution were included in separate waves to avoid nearly-identical emails arriving in the office on the same day (even though the emails are sent to different people). The order of the treatments was randomized so that we have random variation in which institutions received the *Neo-Pronouns* treatment early, for example, and which institutions received it later.

Regarding the selection of email addresses collected, we conducted the following protocol. When there are more than four admission counselors at an institution, we select the first four listed. When there are fewer than four admission counselors we add the next lowest-ranking individual in the office (e.g., Assistant Director).[9]

Our target was to obtain a sample of 2000 contacts at 500 higher education institutions to conduct a within-subject experiment (four per institution). A small number of institutions, though, do not provide all needed information on their web site, or have a small admissions staff. Four of our randomly-selected institutions list their admission counselors but only provide access to an online contact information form. We use the portals for these schools. A total of 18 institutions do have contact information for any admissions office staff. For these, we collect the generic contact email address (e.g., "admissions@xxx.edu"). For an

[7] For example, we received a few replies from individuals who used (she/they) in their signature signaling that they would be comfortable with traditional female pronouns or the singular they pronoun. Other alternatives that could have been used include (ze/zem) and (ne/nem). See pronouns.org for a discussion of pronoun usage and recommended practices.

[8] The pre-registration can be found here: https://doi.org/10.1257/rct.10414-1.0

[9] We try to avoid staff dedicated to transfer students, graduate applications, and international students. Only when necessary do we include contact information from these workers.

additional eight institutions only two email addresses can be obtained, and for six institutions only three addresses could be obtained.

When our list of contact addresses for an institution was incomplete we could not expose it to all four treatments. These institutions always have a generic email address for the office. When the institution only had this one email address, it received two messages spaced out in time. The sampling was done so that the *Control* and *Neo-Pronouns* treatments are sent. If only three addresses could be found and one of them was a generic address, then it would receive two messages so that the institution received all four treatments. If only two addresses could be found and one of them was a generic address, then it would again receive two messages and either the *Male* or *Female* treatment would be omitted (selected randomly). Thus, this protocol ensures that all 500 institutions received the *Control* and *Neo-Pronouns* treatments and that as many as possible received at least one of the other two treatments. Due to incomplete contact information, a small number of institutions do not have one (or both) of the traditional, binary gender treatments. Following this protocol, we end up with a sample of $N = 1957$.

4 Results

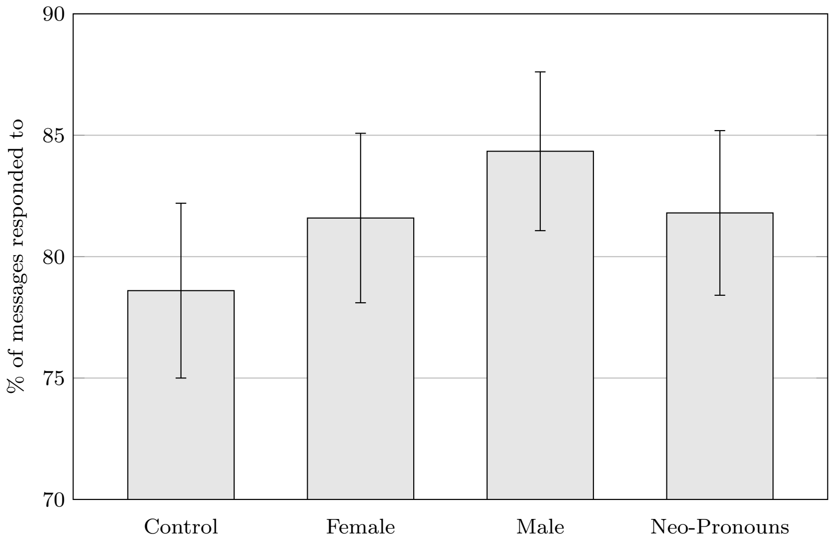
From the 1957 emails sent, we received responses to 81.55% of them.[10] While high, it is potentially surprising that it is much less than 100% as we were posing as prospective students to individuals whose job is to recruit and guide prospective students in their application process.

4.1 Response Rate

We now turn to considering the relationship between the responsiveness and our treatments. Figure 1 graphically depicts the response rates.

[10]While rare, if our message was returned as 'undeliverable', it was recorded as a no-response. In addition, counselors would occasionally turn on their automatic-reply feature when away from their office. These are not recorded as responses either. Usually, the counselor followed up shortly thereafter and that reply was used for our study. If the automated response was never followed up with, then it was recorded as a no-response. Our response rate matches that arising in Hanson [2017] who also sent emails to admission counselors. The response rate there was 72.3% across the treatments.

Figure 1: Response Rate



Each column depicts the proportion of messages in that treatment that were replied to. The 95% confidence intervals are depicted as well. There are $N = 500, 478, 479, 500$ observations, respectively.

Most noteworthy, the response rate for the *Control* is 3.97 percentage points less than the three treatments (combined). To test the statistical significance of these differences Table 1 provides a variety of results.

Table 1: Pairwise Tests

	difference	t-stat
Is the response rate to (he/his) different from (she/her)?	0.0275	1.13
Is the response rate to (xe/xem) different from "traditional" pronouns?	0.0117	0.68
Is the response rate to "traditional" pronouns different from the control?	0.0437	1.88 **
Is the response rate to (xe/xem) different from the control?	0.0320	1.27
Is the response rate to the use of any pronouns different from the control?	0.0397	1.97 **

Results from five difference-in-means t-tests presented; *** 1%, ** 5%, * 10% level of significance.

A number of important observations arise. First, there is not a statistical difference between the use of traditional male pronouns and traditional female pronouns. Thus, moving forward we will combine the two treatments. Second, and quite importantly, the use of neo-pronouns does not experience a different response rate than the use of traditional pronouns. These two distinctions show the smallest discrepancies and t-statistics. Third, though, differing behavior arises in the control. Consistent with Figure 1, the response rate to messages in the *Control* is less, and this difference is significant at the 5% level.

4.2 Regression Analysis

To further verify these observations, we pool the data and estimate a straightforward linear regression. This allows us to add institution fixed effects so that we can explore within-institution variation in responses. Also, while randomized, we can control for order (whether the first message sent to a school is responded to differently than the second, third, or fourth message) and round (whether sent messages earlier in time are responded to differently than later ones). Table 2 presents the results.

Table 2: Regression Results

	(1)	(2)	(3)	(4)
Pronouns	0.0392	0.0417	0.0431	0.0429
	(0.0226) *	(0.0247) *	(0.241) *	(0.0252) *
Neo-Pronouns			-0.0111	-0.0061
			(0.0226)	(0.0259)
University Fixed Effects?	Yes	Yes	Yes	Yes
Round Fixed Effects?	No	Yes	No	Yes
Order Fixed Effects?	No	Yes	No	Yes
R^2	0.371	0.378	0.371	0.378
AIC	942.3	947.5	943.9	949.4

Results from a linear probability model with *Reply* as the dependent variable. The omitted, reference group are the messages sent without the use of pronouns (*Control*). The explanatory variable *Pronouns* is equal to one if the observation arises from any of the three treatments; $N = 1957$. The second and fourth columns include indicator variables for which round the email was sent during (12 rounds) and whether the message was the first, second, third, or fourth message sent to that institution (4 emails). Standard errors clustered at the institution level are presented in parentheses (500 clusters); *** 1%, ** 5%, * 10% level of significance.

The indicator variable *Pronouns* is equal to one if the observation came from any treatment other than *Control*. The results in columns (1) and (2) confirm that those messages which included pronouns received a greater response rate.[11]

For columns (3) and (4) we consider the *Neo-Pronouns* treatment separately. An indicator variable for it is added, which allows us to assess whether the use of neo-pronouns has a differential effect on the response rate. It does not. The marginal effect of a neo-pronoun is highly statistically insignificant.

Taken together, these results suggest that $\nu = 0$ and $\lambda > 0$ from the theoretical model. Thus, Hypothesis 1 receives confirmation but both Hypothesis 2 and $2'$ are contradicted. In other words, the results indicate that there is a preference for progressively-minded students and not a distaste for non-binary individuals.

5 Intensive Margin

While our focus is on whether the admission counselors reply to the request for help, we also explore other outcome variables. While the analysis of the response rate can be thought of as the extensive margin of aid, we ask whether communicating gender identity affects the intensive margin. Rather, are counselors systematically more helpful to some messages?

5.1 Content of the Replies

Measuring the degree of "helpfulness" is not clear cut. We first consider straightforward measurements of the messages received. Specifically, we record the number of hours it took for the counselor to reply, given

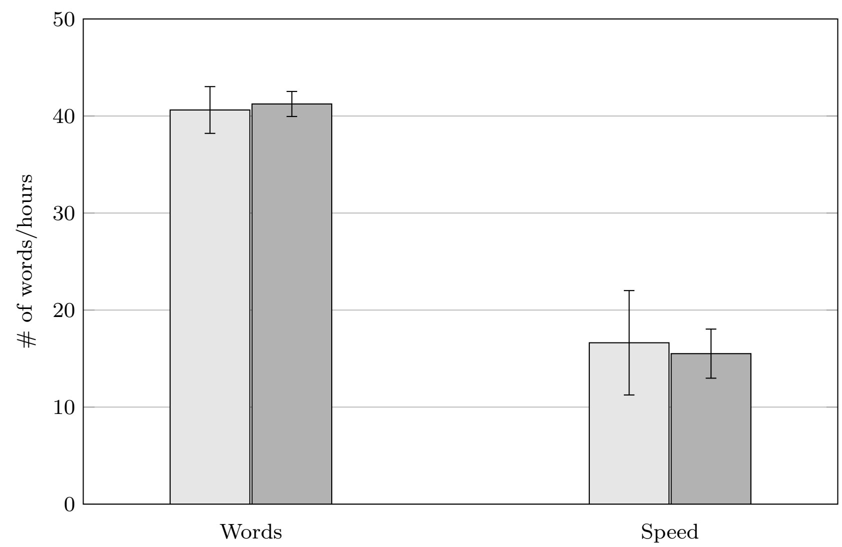
[11] The p-value on the *Pronouns* coefficient is 0.119 in column (2). Without clustering it is 0.065 (and less than 0.05 in the other three columns). Thus, the most conservative hypothesis testing results are presented.

that they replied at all, and the number of words making up the reply. We argue that a greater willingness to help the potential student could lead the counselor to respond quicker and to provide a more detailed response.

The mean response rate is 15.8 hours, with a median of 1.4 hours and a standard deviation of 47.2. The mean word count is 41.1 words, with a median of 37 and a standard deviation of 23.2. Thus, for both measurements there is quite a bit of variation in admission counselor behavior.

We re-conduct the previous analysis, but with these outcomes as dependent variables (i.e., difference-in-means test, nonparametric tests, and the regressions) and fail to find any statistical difference between the treatments. To illustrate, Figure 2 depicts the mean values for both variables when partitioning the sample by whether the observation comes from the *Control* or one of the treatments.

Figure 2: Replies



The left two columns depict the mean number of words used in the reply email. The right two columns depict the mean number of hours until a reply was made (excluding weekends). The light gray columns provide the values for the *Control* and the dark gray columns provide the values for the *Pronouns* treatments. The 95% confidence intervals are depicted as well. There are $N = 400$ & 1207 observations in the two samples, respectively.

This implies, for the response speed, that the decision was whether to respond and the treatments did not affect the eagerness to respond. For the word count, when the admission counselor choose to respond to the message, the quantity of information was unaffected.

It is possible, though, that the quality of information and the tone of the messages differed. To investigate this, we choose some easy-to-measure characteristics of the replies. For one, we record whether the counselor included their own preferred gender pronouns in the message. We also record the use of certain punctuation. Specifically, we create indicator variables for whether the counselor used an exclamation mark and whether they used a question mark in their reply. The first expresses excitement and the second suggests the counselor would like to have a continued communication with the potential student. Further, we recognize when an emoji was included in the text.[12] The presence of such a smile represents a welcoming connection between the applicant and the counselor. Finally, we record whether an email address was included in the message (measured by whether the '@' symbol is present in the reply) and whether a web site was linked (measured by whether 'www' or "http' was present in the reply). Table 3 provides the frequency of each.

Table 3: Text of the Replies

Content of the Reply	Proportion
!	55.01%
?	9.40%
:)	3.61%
Email address	5.72%
Web address	0.81%
Counselor pronouns	27.0%

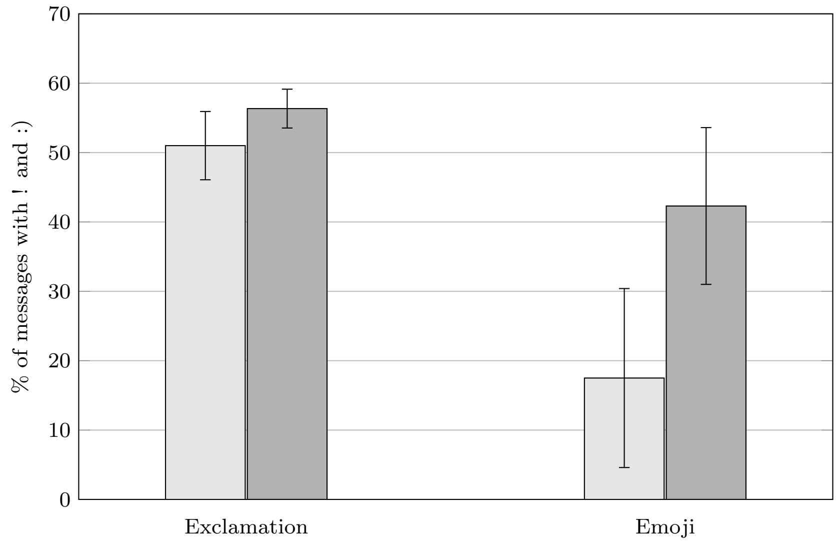
Each is a variable equal to one if the reply includes the item within the text of the reply. We do not consider the number of each in each message, only whether it exists. Thus, each proportion represents the percentage of messages which include each. The indicator variable *Email address* is equal to one if is included in the message, while *Web address* is equal to one if 'www' or 'http' is present.

The use of exclamation marks is quite common. Questions and links to either web sites or email addresses are rather rare. Thus, admission counselors, overall, were not necessarily looking for a continued communication with our potential student.

What is important, though, is whether the frequency of these message characteristics vary by treatment. To this end, we first consider the two positive, friendly variables – the use of exclamation marks and the inclusion of an emoji. Figure 3 compares the *Control* to *Pronouns* for the use of '!' and ':)'. The rate of an emoji/emoticon's use is multiplied by 10 in the figure to have comparable-sized columns.

[12] All emoji images were first converted to the emoticon :) and an indicator variable is created to capture its presence. We do not distinguish between the numerous emoji and emoticon choices selected.

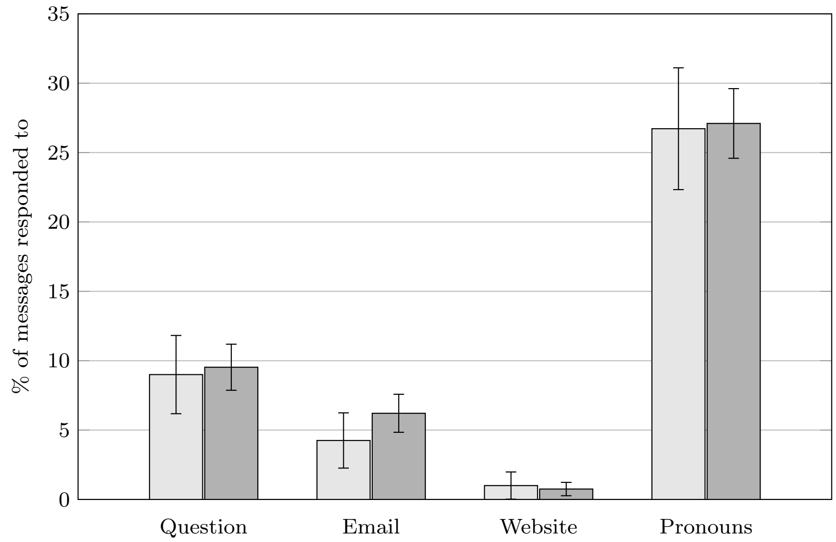
Figure 3: Excitement



The left two columns depict the proportion of replies which included at least one exclamation mark. The right two columns depict the proportion of replies (multiplied by 10 to scale the two) which included an emoji or an emoticon. The light gray columns provide the values for the *Control* and the dark gray columns provide the values for the *Pronouns* treatments. The 95% confidence intervals are depicted as well. There are $N = 400$ & 1207 observations in the two samples, respectively.

When our request for help included preferred pronouns we were more likely to have an exclamation mark included in the reply, conditional on having our message replied to (a 10.5% greater rate). Also, when it included preferred pronouns we were more likely to get an emoji/emoticon included in the reply (a 141.7% increase in the rate). Together, these suggest that admission counselors were more friendly to potential students who are progressively minded. The other measurable characteristics are depicted in Figure 4.

Figure 4: Responsiveness



Each column depicts the proportion of observations in the sample with the indicator variable equal to zero. The light gray columns depict the rate for the *Control* sample and the dark gray depict the rate for those observations in a treatment *Pronouns*. The 95% confidence intervals are depicted as well. There are N = 400/393 & 1207/1203 observations, respectively.

Across these outcome variables, there is not a measurable difference between their prevalence in messages from potential students who include preferred pronouns and from those who do not. Thus, for one, there is no difference in the desire for follow-up communications measured by either the use of a question mark or a link to either an email address or a web site. For the counselors' pronouns, this implies that there was not a reciprocity where counselors added pronouns to their communications when the potential applicant did.

5.2 Topic Modeling

The analysis conducted on the replies received only looks for simple measurable characteristics of the message. It does not necessarily capture the reply's content, but rather focuses on its tone. Here we use tools of topic modeling in an attempt to understand differences in the messages' content.

To do this, we employ a popular topic modeling method known as Latent Dirichlet Allocation (hereafter LDA). First developed by Blei et al. [2003], LDA seeks to uncover the 'themes/topics' by exploring the co-occurrence of words in a text. The logic of LDA is that when an author writes a text, they first choose the topics to be covered. LDA treats this decision as choosing a probability distribution over topics. For this paper, as an illustration, we may have chosen to make it 20% economics, 40% gender, 30% econometrics, and 10% education, but 0% health and 0% labor. Each topic is simply a probability distribution over the dictionary. To continue with the illustrative example, words such as 'hypothesis' and 'regression' may be drawn with a high probability when an econometrics word is to be selected. Words such as 'male' and 'pronoun' may be chosen with a high probability when writing about gender.

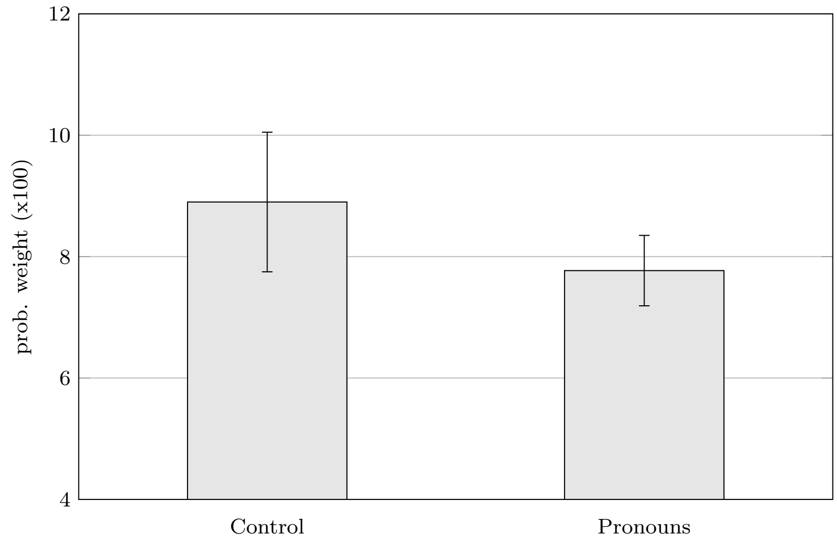
With the presumption of this statistical model of writing LDA seeks to identify the underlying topics that exist in a corpus of documents. It is an unsupervised, machine-learning process that clusters words into topics and topics across documents. The output of the LDA process is a probability distribution for each document across a fixed number of topics. LDA also produces a probability distribution of words for each topic. All the researcher must select is the number of topics they would like to have the documents organized into.

For our analysis, given that the replies' word count are rather small, we choose for LDA to have 10 topics. LDA has been used in a number of contexts in economics research. For example, Dyer et al. [2017] uses LDA to identify trends in 10-K disclosures and Bastani et al. [2019] uses it to gauge investor complaints to the Consumer Financial Protection Bureau. Hansen et al. [2018] evaluates FOMC transcripts using LDA to appreciate how experience on the FOMC and transparency interact. Larsen and Thorsrud [2019] use it to model topics covered in a business newspaper and evaluate the prevalence of the topics in a time series analysis of a measurement of asset prices. McCannon [2020] uses it to evaluate descriptions of wine adding the probability weights created by the LDA process as explanatory variables in a hedonic price regression. Further details and an elaboration on LDA methods can be found in Blei et al. [2003] or Schwarz [2018].

Topic modeling requires us to clean the text. First, we collect the text of each reply received. We collect all text up to, but not including, the counselor's name and their signature line. Second, we engage in standard cleaning. We remove punctuation, hard returns between any lines/paragraphs, and turn all letters to lower cases. We remove stopwords (such as words such as "and", "the", and "is") which do not convey useful information. We adjust past-tense and future-tense words and change plural to singular words. For example, "greetings, becomes "greet". Abbreviations are spelled out so that, for example, "info" becomes "information". Common adverbs are replaced with their adjective form. For example, "correctly" becomes "correct". Without these cleanings, the LDA process would treat each word distinctly. We drop the four names of our fictitious applicants and references to 'college', 'university', or 'state'. The text cleaning reduces the word count to an average of 20.4 words per message down from 41.1.

Following this method, as stated, each reply is assigned a probability distribution over the ten LDA-created topics. For each topic we ask whether there is a distinction in its prevalence between emails sent from a potential student that did not include gender pronouns and messages sent from those with pronouns. One topic stands out. Figure 5 illustrates.

Figure 5: Topic: *Decision*



Each column depicts the average probability weight placed on the topic *Decision*. The 95% confidence intervals are depicted as well. There are $N = 400$ & 1207 observations, respectively. A difference-in-means test has $t = 1.84$ ($p = 0.065$).

This topic is more likely to arise in replies to potential students who do not include their preferred pronouns. Replies to potential students with their preferred pronouns have reply messages focusing on other topics. Figure 4 considers the words that have the highest probability of arising when this topic is selected. The top 15 words and the probability that each is selected is presented. We also consider the rank and probability of that word in the other nine topics. Both the mean and the median are presented.

Table 4: Sample of Words Making Up *Decision*

word	rank	prob	rank - other 9	prob. - other 9	TF-IDF
mail	1	0.0745	57.8/57	0.0092/0.0010	0.0522
address	2	0.0531	12.8/1	0.0841/0.0991	0.0049
receive	3	0.0379	79.4/101	0.0032/0.0008	0.1011
decision	4	0.0331	101/101	0.0007/0.0005	0.1232
send	5	0.0275	59.4/74	0.0063/0.0023	0.0864
admission	6	0.0225	83.3/101	0.0047/0.0005	0.0974
email	7	0.0223	43.0/27.5	0.0126/0.0065	0.0728
use	8	0.0173	30/7	0.0344/0.0476	0.0261
information	9	0.0167	55.4/53	0.0046/0.0014	0.1126
sent	10	0.0163	94/101	0.0010/0.0005	0.1369
letter	11	0.0162	101/101	0.0007/0.0005	0.1449
want	12	0.0153	91/101	0.0015/0.0005	0.1402
time	13	0.0147	48.1/45	0.0038/0.0020	0.1127
most	14	0.0141	101/101	0.0007/0.0005	0.1550
move	15	0.0138	25.8/6	0.0384/0.0426	0.0227

In the fourth and fifth columns the first number is the mean value across the other nine topics and the second number is the median value. Due to a word occasionally falling in the far right tail, we consider only the top 100 words for each topic. If a word does not make this list, it is assigned the rank of 101 and the probability weight equal to the 100^{th} word on that list.

Words frequently arising in this topic are rather distinct from the words used in other topics.[13] This topic focuses on words that factually discuss the application process. The topic discusses the process of sending the decision letter on the application. Thus, we will refer to this topic as *Decision*. To illustrate replies focusing on this topic consider the following three actual messages received.[14]

> [1]
> Hi Morgan, Thanks for your message! The first item we will mail to you is your admission decision. It will be sent electronically and if admitted, also through the mail. Will you have moved by February 1?
>
> [2]
> Hi Morgan, How long until you move? We will send an admissions email and then a physical packet in the mail within about 3 weeks from them so it will depend on that as most communication from WVU will be via email. Warm regards,
>
> [3]
> I would say you should use whichever one you want to get your mail sent to because we'll mail your acceptance letter, scholarship certificate, and financial aid package to that address.

As one can see, these replies are factually discussing the process that decision letters are sent out. The three messages selected are the three with the greatest weight put on the *Decision* topic.[15] Importantly, the first and second messages were sent to potential students who had included the traditional male and female pronouns, respectively. Only the third was from the control sample.[16]

Table 5 verifies that there is not an important difference between the replies to messages in *Control* from the three treatments in the other nine topics.

While the rate the *Control* receives a *Decision*-focused reply is 14.5% more likely than the messages replied to from one of the treatments, the difference across the other nine topics is much smaller.

The topic least likely to be included in replies to messages arising in the *Control* cohort is the final topic listed in Table 5. Among the twenty words most likely to be selected in this topic are 'thank', 'great', 'best', 'please', 'let', and 'know'. These are friendly words. Hence, the *Friendly* topic is more likely to be used in the admission counselor's writing if pronouns are included in the original message.

As with the replies' measurable characteristics, admission counselors who do choose to respond to the request for help are more friendly and enthusiastic to those who communicate preferred pronouns, while they focus only on the factual information of the decision process when pronouns are not included.

[13]The only two words that are also likely to arise in other topics are 'address', 'use' and 'move'. They have very low TF-IDF scores. The TF-IDF measures the value of each word. The score increases when a word is used more, but is decreased when the word is used in more documents. Thus, it measures a word's importance in text classification. Hence, these three words are rather unimportant.

[14]With the text cleaning, these become the following messages. [1] "hi thank message first item mail admission decision sent electronic admit through mail move february 1", [2] "hi long until move send admission email physical packet mail 3 week them depend most communication wvu via email warm regard", and [3] "i say should use one want get mail sent because well mail acceptance letter scholarship certificate financial aid package address".

[15]The weight placed on the *Decision* topic is 0.700, 0.737, and 0.849, respectively.

[16]Further, the replies came from diverse institutions: University of Colorado - Colorado Springs, West Virginia University, and Mary Baldwin College. These are geographically diverse, vary in whether they are public-regional, public-flagship, or private and differ in size, Carnegie classification, and numerous other attributes.

Table 5: All Ten Topics

topic	difference	t-stat	p-value
Topic: *Decision*	0.0113	1.844	0.065
Topic 1	0.0079	1.377	0.169
Topic 2	0.0052	0.663	0.507
Topic 3	0.0041	0.795	0.427
Topic 4	0.0022	0.289	0.773
Topic 5	-0.0012	0.174	0.862
Topic 6	-0.0029	0.413	0.680
Topic 7	-0.0068	1.187	0.236
Topic 8	-0.0085	1.256	0.209
Topic: *Friendly*	-0.0113	1.351	0.177

The topics are ordered in decreasing magnitude of the difference between its prevalence in the *Control* and the three treatments. Positive values for the difference indicates that the *Control* has a greater prevalence for this topic.

6 Using Machine Learning to Predict Replies

A treatment effect exists. A natural question to ask, then, is whether there is treatment effect heterogeneity. There are numerous attributes that one can use to classify an institution of higher education. Rather than select based on our preconceived intuitions as researchers, we rely on a formal approach. We use machine learning techniques to identify which attributes predict responses, and then use these selected attributes to explore heterogeneous treatment effects.

6.1 Random Forests

The machine learning technique we employ is Random Forests [Breiman, 2001]. Briefly, Random Forests builds off of the idea of a classification tree. A classification tree uses attributes to sequentially partition the data with the goal of minimizing prediction errors. Random Forests is a procedure of repeatedly growing these trees. Each uses a random sample of the data and attributes, and classification is done by aggregating predictions across the trees in the forest (a process known as *bagging*). This process identifies which attributes prove to be most important in making accurate classifications. It has been used in numerous machine learning classification problems. For example, Beaulac and Rosenthal [2019] uses it to identify the factors which best predict whether first-year undergraduate students will complete their college degree. Examples in economics include Random Forests being used to predict bank failures [Tanaka et al., 2016] and financial fraud cases [Liu et al., 2015]. Consistently, this approach has been shown to outperform regression-based forecasting using, for example, a logit model.

We collect a number of static attributes for each institution in our sample. Our data comes from the Integrated Postsecondary Education Data System (IPEDS) provided by NCES. We obtain information on each institution's Carnegie classification and numerous indicator variables capturing the characteristics of the institution. They are whether the institution is an HBCU, whether it has a hospital affiliated with it, whether it is a land grant institution, if the school competes in NCAA athletics, and whether it is a public or private institution. We partition the Carnegie classifications to those determined to be doctoral institutions, master's granting, or baccalaureate-focused establishments. An "other" category is used for specialized institutions. Similarly, we categorize the schools by whether they are Catholic institutions, christian (non-Catholic), interdenominational, or do not have a religious affiliation.[17] Further, we create regional indicator variables for whether the institution is located in the west, northeast, south, or midwest/north (as the reference category). Relatedly, we collect information on the institution's urbanization. Specifically, we create four indicator variables distinguishing institutions in cities, suburbs, towns, and rural locations. All attributes except for the region of the country are assessments provided by NCES. We also collect data on the Fall 2021 academic period. We identify the total undergraduate enrollment, along with its gender and racial distribution. Additionally, we record the proportion of the undergraduate population with full-time status and the proportion of the total student body who is in a graduate program. We collect the institution's retention rate and the student-faculty ratio as well. We use the proportion of students receiving any sort of financial aid and the proportion specifically receiving Pell grants to account for the students' financial position. The institution's financial data, such as price charged and university endowment, suffer from under-reporting. Given our modest sample size, we choose not to includes these additional attributes as it would force us to consider only a nonrandom subsample of institutions.[18]

While Random Forests are a black box, one quantifiable outocme is useful for our analysis. Each attribute's importance can be quantified. Briefly, each tree that is grown using a particular attribute is considered. The ability to correctly forecast the outcomes if that attribute was eliminated is measured. Thus, for each variable the number of times that its inclusion has a positive marginal impact on the forest's ability to make the correct prediction is recorded. This is done for each attribute and the number of times it is beneficial, relative to the attribute found most beneficial, is an attribute's *importance*. Thus, each attribute receives an importance score between zero and one, and one of the attributes must have an importance measure equal to one. See Schonlau and Zou [2020] for an elaboration on the Random Forests algorithm and this metric.

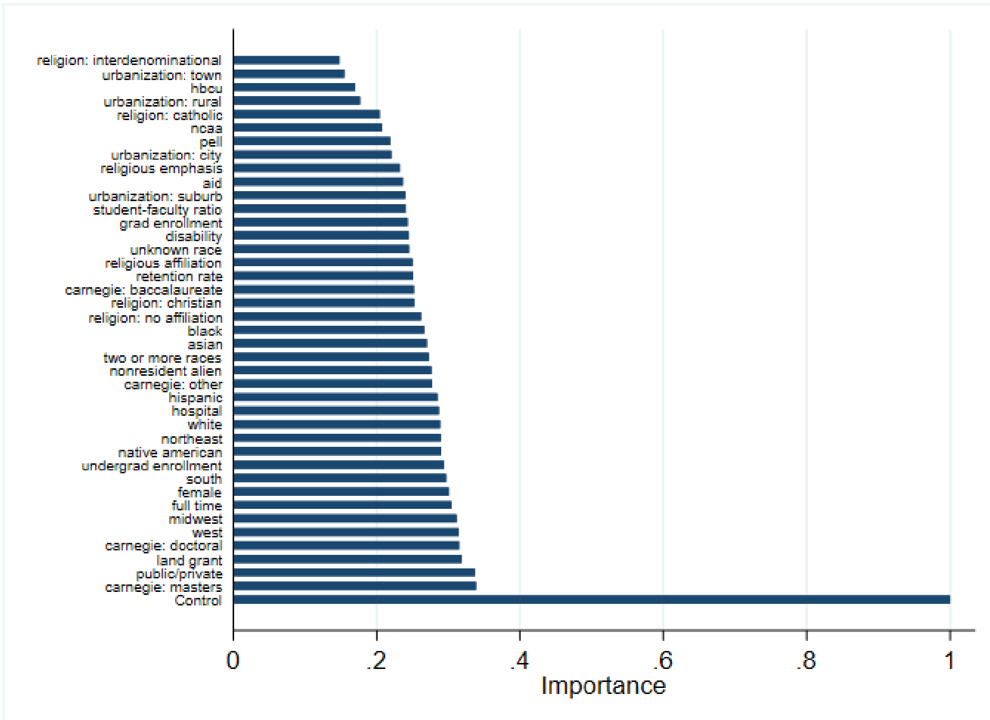
6.2 Heterogeneous Treatment Effects

First, we consider the full data set and include an indicator for being in the control versus being in a treatment which included pronouns. This will rank characteristics by how well they predict a reply. The variable importance is graphically depicted in Figure 6.

[17]In our sample, we do not have a Jewish, Mormon, or Muslim institution. It is important to note that this categorization provided by NCES is based on the institution's origins and does not necessarily capture the intensity of the religious experience on campus. We go further and create our own two categorizations. For them, we review the mission statement and vision for each of the 500 higher education institutions. We record whether they acknowledge a religious history to the institution and whether they specifically signal christian values and theology as part of their core curriculum. We include these as additional attributes.

[18]For example, we lose 8% of the sample if we include price data. The mean undergraduate enrollment for these lost observations is 41.3% lower, with a difference-in-means test having $t = 3.32$ and $p < 0.001$. More worrisome, the response rate is 5.4 percentage points lower for the institutions with missing price ($t = 1.64$; $p = 0.100$). The number of missing observations for endowment is even greater. Therefore, their exclusion could have dramatic impacts on the predictions arising from the machine learning process. Including retention rate, student-faculty ratio, and the financial aid data does require us to drop four institutions who do not report these outcomes.

Figure 6: Variable Importance



The variable importance is depicted for Random Forest algorithm using the full data set.

Clearly, being in a treatment rather than the control is by far the most important attribute to predicting responses to the email inquiries. Also, in the full data set the institutional type (i.e., Carnegie classification), institutional control (i.e., public or private), land grant status, and region of the country are the next most important attributes to forecasting replies.

Our goal, though, is to establish differential effects by treatment. Hence, we start by exploring which attributes correlate with the observed preference for progressively-minded students. To do this, we partition the data into the four treatments and construct a separate Random Forest for each. This allows us to find which attributes are important for classification in the *Control* subsample but are less important for the three treatments.

A number of attributes have an importance greater for *Control* than in any of the treatments. They are whether it is baccalaureate-focused, whether it has a historical religious affiliation, the urban setting of the campus, undergraduate enrollment, retention rate, and the proportion of the students receiving Pell grants. Therefore, we ask whether the preference for progressively-minded students has differential effects across these attributes.

To do this, we modify our previous regression estimation. We interact the treatment indicator (*Pronouns*) with an attribute and add it to the specification. The attribute is included as a control variable as well.[19] A separate specification is estimated for each attribute shown to be more important in the *Control* than any of

[19]Hence, we drop the institution fixed effects. We do continue to cluster the standard errors at the institutional level.

the treatments. From this process we can identify institutional characteristics that have a relatively-greater treatment effect. With this effort, we find that a bias in favor of progressively-minded students is relatively stronger at institutions with[20]:

1. undergraduate student sizes not in the smallest quintile
2. low retention rates
3. large proportion of the student bodies receiving Pell grants
4. a city location.

Hence, institutions with these attributes are more likely to show a preference for students who communicate their preferred pronouns.

6.3 Negative Discrimination

Second, our project is motivated by non-binary gender identity and the admissions process. Thus, we also use this machine learning process to hunt for subsamples of the pool where a negative discrimination against non-binary individuals may exist. Revisiting the variable importance from the Random Forests, numerous attributes show more importance in the *Neo-Pronouns* treatment than the other treatments. This includes the racial makeup of the student body, region of the country, full-time status, student-faculty ratios, financial aid, religious affiliation, HBCU and land grant status, and institutional control. Therefore, we explore potential heterogeneous effects with these attributes.

We follow a similar procedure as before. For each attribute we separately include it in the regression and add an interaction term with the observations being in the *Neo-Pronouns* treatment. The regression also includes an indicator variable for being in one of the traditional gender pronouns treatments. We focus on those attributes which have a statistically significant interaction term. Doing so, we find that a negative discrimination against non-binary individuals, if it is to exist, is most likely to occur at institutions with:

1. low student-faculty ratios
2. small proportion of students receiving financial aid
3. locations in the northeast part of the U.S.

To illustrate this, we create an indicator variable equal to one if an institution demonstrates all of these characteristics; i.e., is in the bottom two quintiles in student-faculty ratio, in the bottom two quintiles for proportion of students receiving financial aid, and if the institution is in the northeast. A total of 14 institutions fall into this category. Looking at the raw data, their response rate in the *Neo-Pronouns* treatment is 21.8% less than their response rate in the traditional gender pronoun treatments and 16.7%

[20]The estimation results are presented in our supplemental appendix. An interaction must be significant at the 10% level to make the list.

less than their response rate in the *Control*. The rest of the institutions in our sample not in this category have response rates in the *Neo-Pronouns* treatment only 1.7% less and 4.1% greater, respectively. Thus, the machine learning process identifies those institutions who disproportionately do not respond to inquiries from non-binary individuals.

7 Conclusion

Experimentally manipulating the inclusion of preferred pronouns in communications, we ask whether there is discrimination in the college admissions process related to gender identity. We find that there is, but (potentially) surprisingly, the aggregate effect coincides with a preference for progressively-minded individuals. Signaling a non-binary status does not have a separate effect. Our results suggest, for example, that concerns about ideological bias on college campuses, which typically focus on faculty and the administration, may also come in part from selection bias in the student population.

We chose to study non-binary gender identity rather than other, closely-related topics, such as transgender status or sexual orientation, due to feasibility. That gender identity is intentionally signaled through the statement of preferred pronouns allows for us to experimentally manipulate this information. Transgender and gender-questioning individuals may very well adopt binary gender identities, which given our design does not allow for separate identification. Similarly, pronoun usage is obviously distinct from sexual orientation. Future work may want to consider avenues to experimentally investigate these complementary issues in a separate audit study.

One potential concern is that our failure to find a separate effect for non-binary individuals could be an artifact of the treatment signal being relatively weak. A number of counselors, for example, may simply not look at the signature line or pay much attention to what is included if they do. We do find a treatment effect overall, so our design is strong enough to identify a consequence to pronoun inclusion broadly. Our design replicates real-world interactions and, therefore, if there is taste-based discrimination against non-binary individuals, we expect it is unlikely to have a marginal impact in the admissions process.

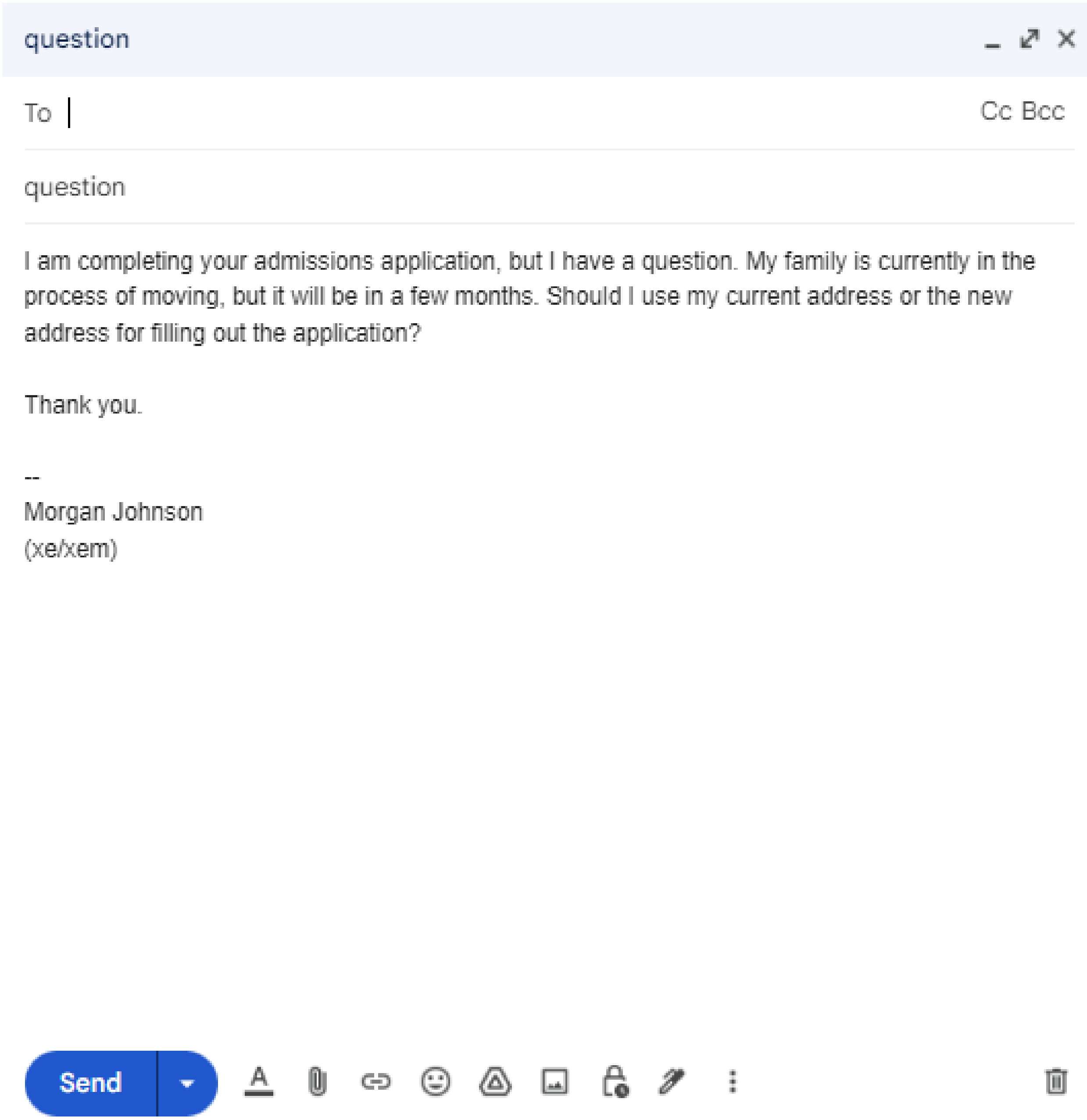
Another issue worth future exploration is the interaction between dimensions to potential discrimination. For example, the names we chose to use in our audit study are not "distinctively Black" names [Fryer Jr. and Levitt, 2004]. It may very well be that the counselors presumed that the applicant was a White-American individual. Transgender and non-binary racial and ethnic minorities could very well experience a stronger discrimination [Button et al., 2020]. Due to a relatively modest subject pool size, we chose to focus on only one dimension to discrimination. Future work, though, may want to investigate this possibility further.

Relatedly, we focused our application to discrimination in higher education motivated, in part, by recent controversies in overt race-based discrimination. Audit studies have searched for discrimination in numerous settings. Prominent examples include labor markets [Bertrand and Mullainathan, 2004], housing [Ewens et al., 2014], and health care [Button et al., 2020]. Along with race, audit studies have explored discrimination based on age, criminal record, and sexual orientation to name a few. Our dual findings of a preference for progressively-minded individuals and a lack of discrimination against non-binary individuals may not arise in differing contexts, and is worth future exploration.

Potentially the greatest leap in the interpretation of our findings is our argument that the inclusion of preferred pronouns signals ideology. This interpretation, though, is essentially untested in our design. It could, alternatively, signal effort put into the application process that is reciprocated by the counselors. Future work may want to directly assess the information presumed when pronouns are included.

8 Appendix

Figure 7: Screenshot



A screenshot of a representative message sent to an admissions counselor. The example depicted would be for the *Neo-Pronouns* treatment.

References

A. Aneja, M. Luca, and O. Reshef. Black Ownership Matters: Does Revealing Race Increase Demand for Minority-Owned Business? *NBER Working Paper 30932*, 2023.

F. Antman and B. Duncan. Incentives to Identify: Racial Identity in the Age of Affirmative Action. *Review of Economics and Discrimination*, 97:710–713, 2015.

T. Bar and A. Zussman. Partisan Grading. *American Economic Journal: Applied Economics*, 4:30–48, 2012.

K. Bastani, H. Namavari, and J. Shaffer. Latent Dirichlet Allocation (LDA) for Topic Modeling of the CFPB Consumer Complaints. *Expert Systems with Applications*, 127:256–271, 2019.

C. Beaulac and J. S. Rosenthal. Predicting University Students' Academic Success and Major Using Random Forests. *Research in Higher Education*, 60:1048–1064, 2019.

M. Bertrand and S. Mullainathan. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94:991–1013, 2004.

D. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3: 993–1022, 2003.

L. Breiman. Random Forests. *Machine Learning*, 45:5–32, 2001.

P. Button, E. Dils, B. Harrell, L. Fumarco, and D. Schwegman. Gender Identity, Race, and Ethnicity Discrimination in Access to Mental Health Care: Preliminary Evidence from a Multi-Wave Audit Field Experiment. *NBER Working Paper 28164*, 2020.

A. Cohen, T. Karelitz, T. Kricheli-Katz, S. Pumpian, and T. Regev. Gender-Neutral Language and Gender Disparities. *NBER Working Paper 31400*, 2023.

J. L. Duarte, J. T. Crawford, C. Stern, J. Haidt, L. Jussim, and P. E. Tetlock. Political Diversity will Improve Social Psychological Science. *Behavioral and Brain Sciences*, 38:e130, 2015.

T. Dyer, M. Lang, and L. Stice-Lawrence. The Evolution of 10-K Textual Disclosures: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, 64:2209–2245, 2017.

M. Ewens, B. Tomlin, and L. C. Wang. Statistical Discrimination or Prejudice? A Large Sample Field Experiment. *Review of Economics and Statistics*, 96:119–134, 2014.

R. G. Fryer Jr. and S. D. Levitt. The Causes and Consequences of Distinctively Black Names. *Quarterly Journal of Economics*, 119:767–805, 2004.

M. Granberg, P. A. Andersson, and A. Ahmed. Hiring Discrimination Against Transgender People: Evidence from a Field Experiment. *Labour Economics*, 65:101860, 2020.

S. Hansen, M. McMahon, and A. Prat. Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. *Quarterly Journal of Economics*, 133:801–870, 2018.

A. Hanson. Do College Admissions Counselors Discriminate? Evidence from a Correspondence-Based Field Experiment. *Economics of Education Review*, 60:86–96, 2017.

V. H. Larsen and L. A. Thorsrud. The Value of News for Economic Development. *Journal of Econometrics*, 210:203–218, 2019.

M. Lawrence and S. Mckendry. *Supporting Transgender and Non-Binary Students and Staff in Further and Higher Education: Practical Advice for Colleges and Universities*. Jessica Kingsley Publishers, London, 2019.

C. Liu, Y. Chan, A. Kazmi, S. Hasnain, and H. Fu. Financial Fraud Detection Model Based on Random Forest. *International Journal of Economics and Finance*, 7:178–188, 2015.

R. J. MacCoun and S. Paletz. Citizens' Perceptions of Ideological Bias in Research on Public Policy Controversies. *Political Psychology*, 30:43–65, 2009.

B. C. McCannon. Wine Descriptions Provide Information. *Journal of Wine Economics*, 15:203–218, 2020.

M. Schonlau and R. Y. Zou. The Random Forest Algorithm for Statistical Learning. *Stata Journal*, 20:3–29, 2020.

C. Schwarz. ldagibbs: A Command for Topic Modeling in Stata Using Latent Dirichlet Allocation. *Stata Journal*, 18:101–117, 2018.

R. Stewart and C. Uggen. Criminal Records and College Admissions: A Modified Experimental Audit. *Criminology*, 58:156–188, 2020.

K. Tanaka, T. Kinkyo, and S. Hamori. Random Forests-Based Early Warning System for Bank Failures. *Economics Letters*, 148:118–121, 2016.

S. Waite. Should I Stay or Should I Go? Employment Discrimination and Workplace Harassment against Transgender and Other Minority Employees in Canada's Federal Public Service. *Journal of Homosexuality*, 68:1833–1859, 2021.