1 Introduction

Undergraduate borrowing and the subsequent required repayments are some of the largest determinants of post-baccalaureate educational investment. Debt itself can act as a disincentive for further borrowing and limit access to external avenues of funding for graduate education. In turn, these consequences affect if and when a student may invest in further education and pursue graduate studies. Furthermore, investments in education present trade-offs with consumption and savings; a rational consumer will only invest if the returns on this investment in the future are high enough to smooth consumption over the lifetime. If we take graduate education to improve labor force efficiency, any deterrence to higher education has potential negative feedback effects on the larger economy. Thus, we are interested in determining whether undergraduate borrowing has an effect on graduate degree attainment? If so, to what degree?

This paper proceeds as follows: The remainder of this section presents some related literature and the motivation behind this work. Section 2 gives an overview of the data used for analysis. Section 3 describes the concept of double machine learning and how it is used to estimate the average treatment effects for undergraduate borrowing, and Section 4 summarizes the results obtained. The paper concludes with a discussion of its broader impacts and limitations in Section 5.

1.1 Related Literature

Over the last several decades, a variety of studies have directly attempted to estimate the effect of undergraduate borrowing on graduate school enrollment through the use of statistical methods. Millett (2003) uses a logistic regression model and results obtained from the 1993 Baccalaureate and Beyond Longitudinal Survey to conclude that students with at least $5,000 of undergraduate debt are significantly less likely to enroll in graduate school. Malcom and Dowd (2012) use a propensity score matching technique to find negative effects of borrowing on graduate school enrollment for every racial-ethnic group analyzed in the 2003 National Survey of Recent College Graduates. Witteveen (2023) uses a two-stage least squares Lewbel method and the 2017 National Survey of College Graduates and finds that while grants bolster graduate school enrollment, undergraduate loans deter students from graduate studies.

On the other hand, Kim and Eyermann (2006) find a positive effect of borrowing on graduate school enrollment in the context of the Higher Education Amendments of 1992 through the use of a logistic model, and Chen and Bahr (2020) find that undergraduate debt has very minimal effects on graduate school attendance through the use of a stratification-based marginal mean weighting method on the 2008-2012 Baccalaureate and Beyond Survey. Notably, the literature on this subject is inconclusive, and most of the works discussed rely only on statistical methods. This background provides an opportunity to obtain results through the use of machine learning techniques.

1.2 Motivation

Zhao (2024) establishes statistically significant results on the negative effects of loans on graduate school enrollment by estimating several logistic regression specifications of graduate school enrollment on undergraduate loans, along with other observed demographic and education-related covariates. However, these results are established on the assumption that the amount of loans and covariates all relate linearly to the outcome variable, which may not be the case in reality. Furthermore, the results in Zhao (2024) only provide evidence for a correlation between undergraduate borrowing and graduate school enrollment; the results cannot be interpreted as a causal effect on post-baccalaureate schooling decisions.

The use of machine learning presents an opportunity to relax the assumptions made by Zhao (2024) and the related works discussed previously, while achieving results which estimate the average treatment effects for borrowing.

2 Data: The National Survey of College Graduates

This paper reuses the pre-processed data from Zhao (2024) which aggregates portions of the 2013, 2015, 2017, 2019, and 2021 iterations of the National Survey of College Graduates (NSCG) (National Center for Science and Engineering Statistics 2024) and converts any relevant categorical variables into binary indicator variables. The survey itself focuses on college graduates who hold at least a bachelor's degree and asks about demographic background, schooling choices, and occupational outcomes.

The only modification to the dataset is the renaming of the columns to better represent the information in each column; no new observations are added, and no existing observations are dropped, so the dataset contains 119,195 observations. Although the NSCG only provides loan information in the form of $10,000 categories, it is the best source of data which includes information about individuals' backgrounds and is publicly available. Other sources of data such as the National Postsecondary Student Aid Study exist, but are not available to the wider public without a restricted use license.

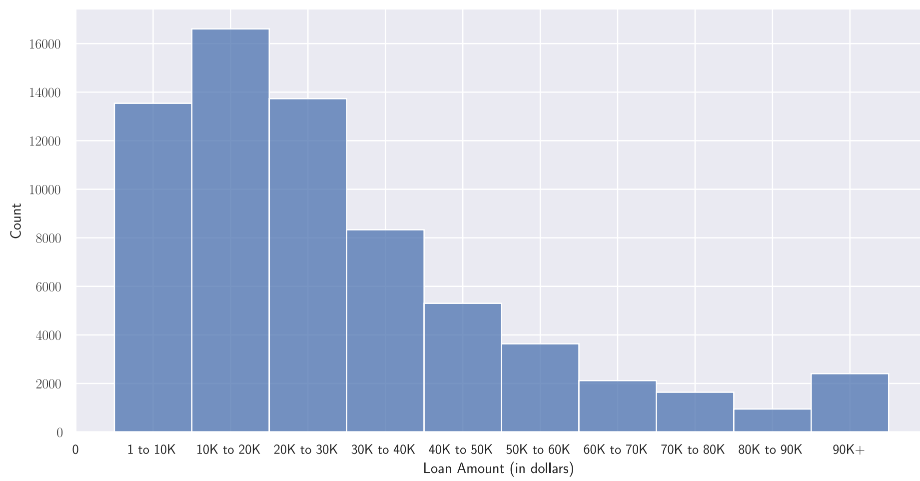
Summary statistics for a selection of the variables in the dataset are supplied in the table below:

Table 1: Observations by select demographic characteristics.

Variable	Num. of. Obs	Share
Gender		
Female	58,713	49.258%
Male	60,482	50.742%
Race		
Asian only	13,165	11.045%
American Indian / Alaska Native only	861	0.722%
Black only	11,436	9.594%
White only	87,633	73.521%
Native Hawaiian / Other Pacific Islander only	595	0.499%
Multiple Race	5,505	4.618%
Hispanic		
No	104,518	87.687%
Yes	14,677	12.313%
Highest Parental Education Level		
Less than high school completed	6,245	5.231%
High school diploma or equivalent	19,377	16.257%
Some college, vocational, or trade school	25,015	20.987%
Bachelor's degree	31,017	26.022%
Master's degree	23,061	19.347%
Professional degree	6,794	5.700%
Doctorate	6,976	5.853%
Not applicable	710	0.596%

Unsurprisingly, there is a significant right skew to the amount that individuals borrow, among those who require undergraduate loans. There are several explanations for this trend. Some US institutions provide financial aid in the form of grants which do not require repayment, reducing the sticker price of tuition. Public institutions are also significantly cheaper for those who have residences within the state. Finally, the federal government also caps the total amount that undergraduates can borrow to $57,500, although some students seek private sources for loans.

Figure 1: Histogram of non-zero loan categories.



Almost two-thirds of the individuals in this dataset only hold a bachelor's degree. While the amount of graduate degree holders has generally risen over time, preferences for schooling or financial barriers (including cost and perceived degree returns) remain as limitations to post-baccalaureate enrollment.

3 Methodology

We seek to estimate the average treatment effect (ATE) of undergraduate borrowing on graduate degree attainment. In an ideal experiment, we would evaluate the differences in decision-making for the exact same individual in two scenarios: one where the individual did not borrow as a means to finance his bachelor's degree, and one where the individual took on some level of debt to finance his bachelor's degree. However, it is impossible to observe both scenarios simultaneously for one individual, so we typically estimate causal parameters through randomized trials. In this context, we would blindly assign the treatment, undergraduate debt, to one group of students who have preferences for continued schooling, and assign the control to another. As the control group represents students without the need to borrow, the experiment might supplement students' finances or provide grants. After the conclusion of the bachelor's program, we would be able to directly compute the average treatment effects based on the students' decisions to enroll or not to enroll in graduate school.

A randomized trial as described is infeasible; the best data available is observational. Given an infinite amount of observational data and computational power, we would most likely be able to construct a sufficiently large dataset to account for all measurable confounding variables, reducing the effects of measurement error or biased results. We would also be able to fit more complex machine learning algorithms, such as neural networks, to use in estimating the average treatment effect with methods such as double machine learning (DML).

3.1 Double Machine Learning

Realistically, we must compromise in our methodology to find a representative estimation of the ATE under the restrictions of the NSCG and the limitations of computational power. Double machine learning, introduced by Chernozhukov et al. (2018), is a technique used to estimate the average effects of a treatment or policy. At a high level, this technique relies on two regressions, one to estimate the treatment and one to estimate the outcome. The first regression is akin to estimating a propensity score for treatment, conditional on a set of covariates. Estimation is conducted using any machine learning method and is typically context dependent. Finally, treatment effects are calculated using the residuals from each regression. DML addresses biases by using a score function which is adjusted such that its expectation is zero, indicating orthogonality between the residuals and the confounders. Known as the Neyman orthogonality condition, this expectation ensures that the correlation between the estimation of the nuisance parameters and the residuals is minimized, refining the estimate of the average treatment effect.

Double machine learning relies on a partially linear model which takes the following form:

$$Y = D\theta_0 + g_0(X) + U,$$
$$D = m_0(X) + V,$$

with the assumptions that $\mathbb{E}[U \mid D, X] = 0$ and $\mathbb{E}[V \mid X] = 0$. Y represents the outcome variable, or the binary variable of graduate degree attainment in this paper. D represents the treatment variable, and θ_0 is the corresponding causal effect of the treatment. The treatment is discussed in depth in the following sections. Finally, X represents the set of covariates which we believe to have both an impact on the treatment itself and the outcome. In context, this includes age, gender, race, ethnicity, parents' education, bachelor's year, the public or private status of an institution, and the field of study.

The variable of primary interest in this paper is undergraduate borrowing; we consider this to be the treatment or policy for which we want to evaluate the effects of. We first establish a baseline result by estimating the partially linear model on a single binary indicator variable equal to one if an individual took on loans to finance a bachelor's degree. If the results of the previous literature are largely representative of the actual effects of undergraduate borrowing, we hypothesize the treatment effect to be negative and the magnitude of the effect to be an average effect of any amount of borrowing. We then replace the indicator variable with categorical variables representing loan bins of $10,000 and repeat the analysis. The average treatment effect on each category is also expected to be negative, although the magnitude of the effect should differ as individuals are likely to make different decisions based on the constraints imposed by the amount of debt.

3.1.1 Alternative Model

The `DoubleML` package also provides the ability to use an interactive regression model specified as follows:

$$Y = g_0(D, X) + U,$$
$$D = m_0(X) + V,$$

with the assumptions that $\mathbb{E}[U \mid X, D] = 0$ and $\mathbb{E}[V \mid X] = 0$. The average treatment effect in this model is computed as:

$$\theta_0 = \mathbb{E}[g_0(1, X) - g_0(0, X)].$$

In this model, the treatment D must be a binary variable; so, we test only the binary loan indicator variable. We obtain estimations for the average treatment effect on this model to use as comparison to the results obtained from the partially linear model. If the results are generally similar, we have reason to believe that our estimates are robust. For all results obtained, we test a variety of learner combinations, and we tune each learner's hyperparameters by selecting multiple candidate values and evaluating the resulting root mean-squared errors of the nuisance functions.

4 Results

4.1 Loan Indicator Variable

Using the binary loan indicator variable, our results suggest that there is a statistically significant negative treatment effect on graduate degree attainment. In particular, an individual with any non-zero amount of loans is on average 4.61% less likely to attend a graduate program.

Table 2: Estimated average treatment effect for the loan indicator using gradient boosting.

Gradient Boosting	ATE	Std. Error	t	[Conf. Interval.]	
Loan Indicator	-0.046	0.003	-16.107	-0.052	-0.041

Because this result does not differentiate between the amount of loans taken by an individual, we can only interpret this result as an overall average. It is possible that the effects vary depending on the level of borrowing; we explore this possibility in the following sections.

4.2 Loan Categories

To estimate the average treatment effects for each loan category, we choose four different learners to serve as comparison: (1) a combination of lasso and logistic regression, (2) decision tree learners, (3) random forest learners, and (4) extreme gradient boosting learners. The results are summarized in the table below:

Table 3: Estimated average treatment effects for the loan categories in the partially linear model across four learners.

Linear	ATE	Std. Error	t	[Conf. Interval.]	
1 to 10K	0.001	0.004	0.256	-0.008	0.010
10K to 20K	-0.012	0.004	-3.093	-0.020	-0.005
20K to 30K	-0.038	0.004	-8.695	-0.046	-0.029
30K to 40K	-0.054	0.005	-10.073	-0.065	-0.044
40K to 50K	-0.065	0.007	-9.521	-0.079	-0.052
50K to 60K	-0.049	0.008	-6.211	-0.065	-0.034
60K to 70K	-0.058	0.010	-5.560	-0.078	-0.037
70K to 80K	-0.047	0.012	-3.998	-0.070	-0.024
80K to 90K	-0.033	0.015	-2.152	-0.063	-0.003
90K+	0.006	0.010	0.601	-0.013	0.025

Decision Tree	ATE	Std. Error	t	[Conf. Interval.]	
1 to 10K	0.001	0.004	0.126	-0.008	0.009
10K to 20K	-0.009	0.004	-2.101	-0.017	-0.001
20K to 30K	-0.036	0.004	-8.180	-0.044	-0.027
30K to 40K	-0.054	0.005	-9.933	-0.065	-0.043
40K to 50K	-0.051	0.007	-7.628	-0.064	-0.038
50K to 60K	-0.044	0.008	-5.435	-0.060	-0.028
60K to 70K	-0.052	0.010	-5.004	-0.073	-0.032
70K to 80K	-0.048	0.012	-4.058	-0.072	-0.025
80K to 90K	-0.037	0.016	-2.354	-0.068	-0.006
90K+	0.004	0.010	0.354	-0.016	0.023

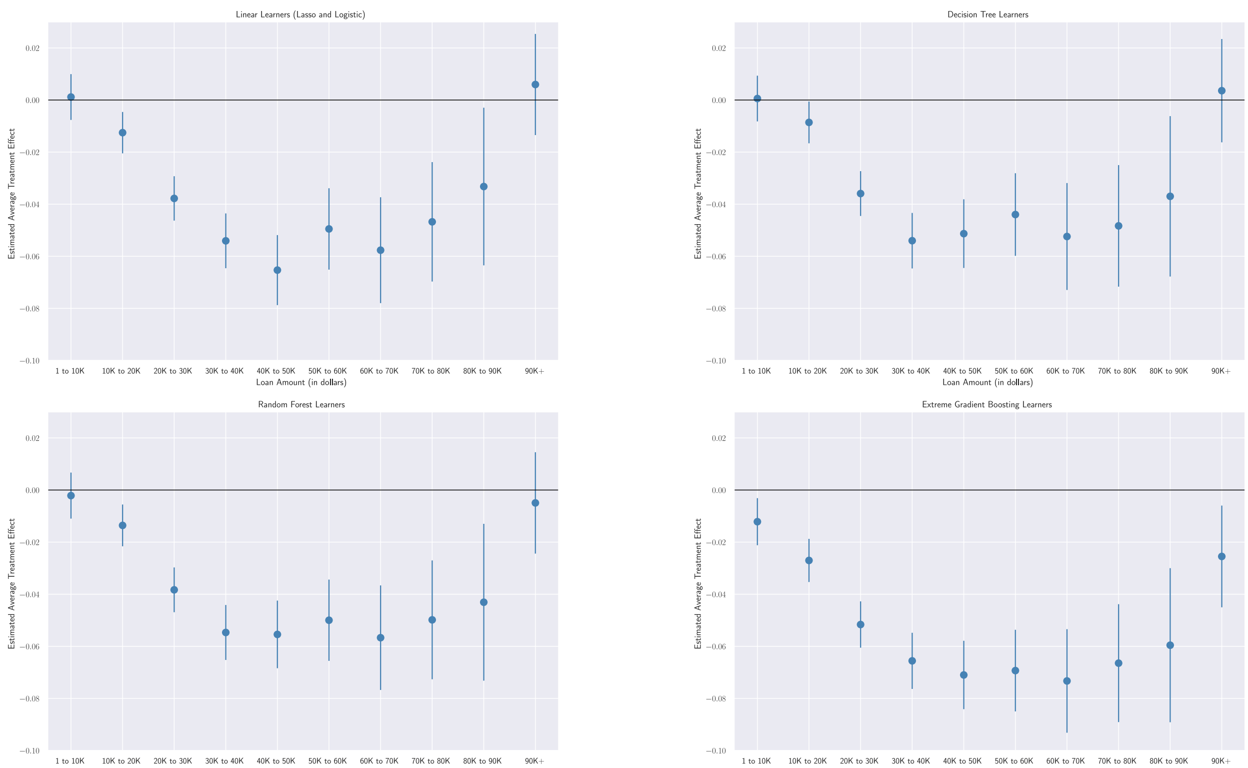
Rand. Forest	ATE	Std. Error	t	[Conf. Interval.]	
1 to 10K	-0.002	0.005	-0.474	-0.011	0.007
10K to 20K	-0.014	0.004	-3.319	-0.022	-0.006
20K to 30K	-0.038	0.004	-8.741	-0.047	-0.030
30K to 40K	-0.055	0.005	-10.159	-0.065	-0.044
40K to 50K	-0.055	0.007	-8.371	-0.068	-0.042
50K to 60K	-0.050	0.008	-6.282	-0.066	-0.034
60K to 70K	-0.057	0.010	-5.537	-0.077	-0.037
70K to 80K	-0.050	0.012	-4.283	-0.073	-0.027
80K to 90K	-0.043	0.015	-2.805	-0.073	-0.013
90K+	-0.005	0.010	-0.498	-0.024	0.015

Extreme GB	ATE	Std. Error	t	[Conf. Interval.]	
1 to 10K	-0.012	0.005	-2.637	-0.021	-0.003
10K to 20K	-0.027	0.004	-6.395	-0.035	-0.019
20K to 30K	-0.052	0.005	-11.375	-0.060	-0.043
30K to 40K	-0.066	0.006	-11.924	-0.076	-0.055
40K to 50K	-0.071	0.007	-10.601	-0.084	-0.058
50K to 60K	-0.069	0.008	-8.673	-0.085	-0.054
60K to 70K	-0.073	0.010	-7.230	-0.093	-0.053
70K to 80K	-0.066	0.012	-5.756	-0.089	-0.044
80K to 90K	-0.060	0.015	-3.948	-0.089	-0.030
90K+	-0.025	0.010	-2.558	-0.045	-0.006

Notably, the results obtained in each of these results are larger in magnitude than the results obtained in Zhao (2024), but the curvature formed by the categories follows closely to the marginal effects in Zhao (2024), suggesting non-linearity in the effects of undergraduate borrowing which become apparent at higher levels. Some students may put more weight on future potential earnings, and thus be more willing to take on loans to invest in education.

Interestingly, the magnitude of the average treatment effect is the largest for the extreme gradient boosting learners. It is also the only learner that predicts a significant negative effect for the highest loan group which represents borrowing at $90,000 or more. Generally, we observe confidence intervals containing 0 for this category as a consequence of the high levels of noise within this group.

Figure 2: Comparison of estimated treatment effects for the loan categories across four learners.



Notes: The top left uses a combination of lasso and logistic regression learners; the top right uses decision tree learners; the bottom left uses random forest learners; and the bottom right uses extreme gradient boosting learners.

4.3 Interactive Regression Model

We use the same four sets of learners to estimate the average treatment effect on the loan indicator variable with the interactive regression model. Across all learners except the decision tree set, we find significantly similar estimates. In comparison to the estimate obtained on the partially linear model, there is virtually no statistically significant difference; the confidence intervals all overlap.

Table 4: Estimated average treatment effects for the loan indicator in the interactive regression model across four learners.

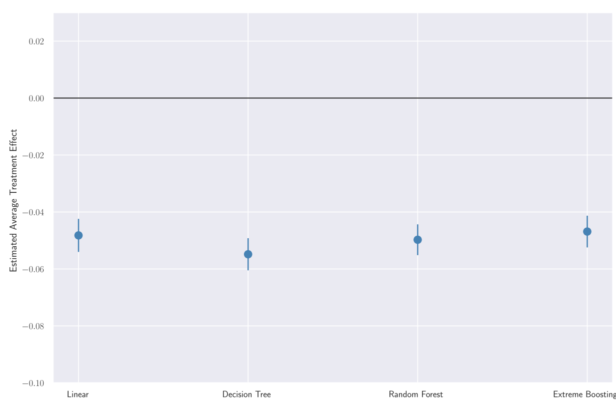
Linear	ATE	Std. Error	t	[Conf.	Interval.]
Loan Ind.	-0.048	0.003	-16.288	-0.054	-0.042

Decision Tree	ATE	Std. Error	t	[Conf.	Interval.]
Loan Ind.	-0.055	0.003	-19.069	-0.060	-0.049

Rand. Forest	ATE	Std. Error	t	[Conf.	Interval.]
Loan Ind.	-0.050	0.003	-18.028	-0.055	-0.044

Extreme GB	ATE	Std. Error	t	[Conf.	Interval.]
Loan Ind.	-0.047	0.003	-16.493	-0.052	-0.041

Figure 3: Estimated average treatment effect for the loan indicator in the interactive regression model across four learners.



Across all four learners, decision tree produces results which differ slightly when compared to the other three. Decision tree is generally more susceptible to overfitting and high variance, which may explain this observation. However, all estimates overlap in their confidence intervals, and are thus generally accepted as accurate representations of the average treatment effect in this model.

5 Discussion and Concluding Remarks

5.1 Limitations

The limitations described in the motivating paper carry forward to this project. These limitations are generally attributed to the data provided by the NSCG. As mentioned previously, undergraduate loan information is only available in $10,000 categories; if we had the exact amount that each individual borrowed, we might be able to obtain better estimates of the treatment effect and better capture the non-linearity trends in larger loan amounts. Furthermore, we can only select confounders which are present in the survey. Parents' education is a fair proxy for family wealth, but it probably does not accurately account for outliers within each education group. Information on household income would probably yield better estimates. Additional confounders not included in the survey such as variables which represent business cycle factors at the time of graduation, expected average income for different fields of study at both the undergraduate and graduate level, and potential graduate debt are all likely to affect the treatment and the outcome.

Due to computation limits, we are also unable to tune the relevant hyperparameters for each classifier as much as might have preferred to. For example, searching over a grid of two parameters with three options each, for two gradient boosting learners required to perform double machine learning, takes more than sixty minutes to select the optimal hyperparameters from the grids. To balance the trade-offs between model performance and computation time, we try a few combinations of hyperparameters for each of the models used in the results section and include the best parameters tested in the accompanying code.

5.2 Conclusion

The results obtained from the method of double machine learning appear to agree with the motivating paper, Zhao (2024), and most other related works relying on statistical methods which find debt to be a deterrence to graduate studies. We observe a negative average treatment effect for undergraduate borrowing in the partially linear model for both the loan indicator variable and the loan categories. We observe some differences in the effects across different loan categories, although the exact non-linear form is indeterminable. The interactive regression model yields similar results, supporting the robustness of the conclusions presented in this paper. Future work might address the limitations discussed previously by using other sources of data and working on machines with better computational power.

References

Chen, Rong and Peter Riley Bahr (Sept. 2020). “How Does Undergraduate Debt Affect Graduate School Application and Enrollment?” In: *Research in Higher Education* 62.4, pp. 528–555. DOI: 10.1007/s11162-020-09610-y.

Chernozhukov, Victor et al. (Feb. 1, 2018). “Double/debiased machine learning for treatment and structural parameters”. In: *The Econometrics Journal* 21.1, pp. C1–C68. DOI: 10.1111/ectj.12097.

Kim, Dongbin and Therese S. Eyermann (July 1, 2006). “Undergraduate Borrowing and Its Effects on Plans to Attend Graduate School Prior to and After the 1992 Higher Education Act Amendments”. In: *Journal of Student Financial Aid* 36.2. ISSN: 0884-9153. DOI: 10.55504/0884-9153.1060.

Malcom, Lindsey and Alicia Dowd (Dec. 2012). “The Impact of Undergraduate Debt on the Graduate School Enrollment of STEM Baccalaureates”. In: *Review of Higher Education* 35, pp. 265–305. DOI: 10.1353/rhe.2012.0007.

Millett, Catherine M. (2003). “How Undergraduate Loan Debt Affects Application and Enrollment in Graduate or First Professional School”. In: *The Journal of Higher Education* 74.4, pp. 386–427. ISSN: 1538-4640. DOI: 10.1353/jhe.2003.0030.

National Center for Science and Engineering Statistics (2024). *Scientists and Engineers Statistical Data System (SESTAT)*. National Science Foundation. URL: http://www.nsf.gov/statistics/sestat/.

Witteveen, Dirk (Mar. 2023). “The Effects of Undergraduate Financing on Advanced Degree Attainment”. In: *Social Forces* 101.3, pp. 1258–1287. DOI: 10.1093/sf/soac044.

Zhao, Amy (Apr. 2024). “Assessing the Effects of Undergraduate Borrowing on Graduate Degree Attainment and Potential Deferral”.