Appendices to Medicaid, Earnings, and Heterogeneous Treatment Effects: Evidence from the Oregon Health Insurance Experiment

A1. Discussion on Theoretical Predictions of Static Labor Supply Theory

This section offers an extended discussion on the theoretical predictions underlying the descriptive value of using distributional analysis. Figure I from Yelowitz (1995) graphically demonstrates the labor supply relevance of Medicaid to low−income households. Medicaid represents an implicit transfer to participating households, equal to the monetized value of Medicaid's health benefits and insurance against health care debt. This is visually represented in Figure I by the parallel shift of the budget constraint. While the benefit of Medicaid does not vary by levels of income for participating households, the income cutoff at Point G demonstrates that Medicaid eligibility is directly determined by household income.

The change in optima under Medicaid for different vintages of workers demonstrates the model's prediction of behavioral heterogeneity. A worker at the point A non−negativity bound cannot further reduce his labor supply. He achieves higher utility under Medicaid at B but exhibits no behavioral response. Alternatively, the constraint does not bind workers at C. They re−optimize at D, demonstrating the strictly negative local prediction below the eligibility threshold for employed compliers with non−zero earnings. The eligibility threshold introduces a budget set non−convexity at F. Those individuals optimizing in this region have weaker preferences for leisure than the previous agents and bunch at the "Medicaid notch" − which just ensures program eligibility. The agent moving to F from ineligibility at E offers an example of bunching due to Medicaid's benefit cliff.

The budget set non−convexity at F dominates a significant range of pre−treatment optima. Nonetheless, the notch−dominated segment of the budget constraint should be bounded. This occurs at a level of pre−treatment earnings where one is just indifferent or prefers her preexisting labor−leisure bundle to the notch. The agent at point J demonstrates this case. There is no incentive to reduce labor supply as relocation is utility diminishing. These individuals represent the other reason the model's global predictions are defined as weakly negative, alongside workers constrained by the zero bound.

Summarizing all four cases, static labor supply theory suggests individuals respond heterogeneously to Medicaid. Agents with zero labor supply or at relatively high levels of earnings should not alter their behavior. Those with pre−treatment earnings near or within the eligibility threshold decrease labor supply. Provided that Figure I is a relevant frame for explaining the causal effect of Medicaid on labor supply, empirical evidence should align with its predictions. Furthermore, these predictions suggest that the sample composition of worker pre−treatment earnings will have implications for average treatment effect estimates. A sample composed of workers at point A pre−treatment should lead to a different estimate of Medicaid's average impact on earnings than a sample dominated by point C or E workers.

A2. Diagnostic Testing for Baseline Distributional Equivalence

Estimating average treatment effects within an experimental setting requires testing that pre−treatment sample means are not statistically different. This verifies successful random assignment and that estimates can be interpreted as causal effects. This is the identifying assumption used in Baicker, et al. and Finkelstein, et al. The bar for asserting successful randomization is higher in the distributional context considered in this paper. It is necessary to prove that the baseline earnings distributions are statistically equivalent. Both strategies follow from the underlying logic of constructing valid comparison groups for reliable inference. If the distributions are different, estimated QTE's may erroneously report underlying differences (or the lack thereof) as the true treatment effect. Both empirical probability density functions and diagnostic testing can verify pre−treatment distributional equivalence.

A2.1 Randomized Treatment Identification – Testing for Distributional Equivalence

Figure II displays kernel density functions of earnings plotted by the number of OHIE eligible members in the responding household. Conditional earnings distributions are used as winning the lottery correlates with the number of OHIE eligible members in a household. The sample consists of treated and untreated households responding to both OHIE baseline and one−year surveys. K−density functions for households with only one OHIE eligible member ("single eligible households") in Panel A appear equivalent based on visual inspection. The treated group empirical density function almost completely overlaps the control PDF. Distributional equivalence for the double eligible households in Panel B is less clear, with the treated group's empirical density function shifted slightly to the right of the control density.

Although the visual evidence offered by empirical density functions is useful, it cannot offer statistical evidence of distributional equivalence. The paper utilizes two diagnostic methods to formally test the equality of the conditional earnings distributions by treatment status. Results are found in Panel A of Table A1. Kolmogorov−Smirnov (K−S) test results are found in columns 1 and 2.[1] As seen in column 2, the K−S test fails to reject the null hypothesis that the baseline distributions are equivalent at the 10% level for single eligible households. Moving to columns 3 and 4, the Wilcoxon−Mann−Whitney rank sum test offers the same conclusion of treated−control distributional equivalence in the single eligible

[1] Kolmogorov−Smirnov testing evaluates whether two distributions are the same through a two−step process. It first identifies the greatest difference in distance between two distributions across all quantiles. Having identified where the vertical distance is greatest, it then tests whether this difference is statistically significant. More background on the K−S test can be found in Imbens and Woolridge (2009).

sample.[2] The results are mixed for the double eligible households. The Wilcoxon test marginally fails to reject distributional equivalence for the double eligible households at the 10% level. However, the K−S test rejects distributional equivalence at the 10% level.

A2.2 Evidence for Differential Attrition or Dissimilar Baseline Distributions

While not a concern for the single eligible OHIE households, the evidence from Figure II and Table A1 suggests that the earning distributions for double eligible households (responding to the one year survey data) are dissimilar at baseline. Failure to assert distributional equivalence could be related to two separate concerns: errors in the random assignment of treatment or differential attrition out of the survey. The first implies that the baseline distributions of earnings are themselves dissimilar. Hypothetically, this outcome could be consistent with the arguments for successful randomization made by Finkelstein et al. The earnings distributions for the treated and control groups could be statistically different yet share a common mean. To test this hypothetical, one can conduct diagnostics testing on all baseline survey respondents. The absence of evidence for distributional differences in this sample would suggest attrition as the real source of concern. In this case, provided the baseline distributions are found to be equivalent, restricting the sample to the subset of survey respondents participating in both survey waves produces the distributional in−equivalence. Non−random patterns of survey non−response change the shape of the treated or control earnings distribution.

Figure III contains kernel density approximations of both treatment and control groups using just the baseline survey sample. Density functions of earnings are plotted for both samples

[2] Wilcoxon−Mann−Whitney rank sum tests combine all observations from both samples, ranks them by magnitude, and computes a sum of these ranks. It then tests the significance of the difference between both rank sums. More information can be found in Kroenker (2005).

by the number of OHIE eligible members in the household. Looking at both density functions, it appears that the earnings distributions are practically the same.[3] This is especially true for the single eligible households, where the treated density essentially overlaps the control. More importantly, differences in the double eligible densities are still quite small. To the extent gaps exist at 70% FPL and between 125−200% FPL, they are consistent with lottery winners earning slightly more. Both graphs align with successful randomization but statistical testing is still needed to verify the distributions as equivalent.

Similar to before, both the K−S and Wilcoxon diagnostic tests are used to evaluate the distributional equivalence of conditional earnings distributions for all baseline survey respondents. Results from both tests are reported in Panel B of Table AI. After accounting for differences in eligible members, the K−S test fails to reject that both baseline distributions are the same. A similar result is confirmed by the Wilcoxon−Mann−Whitney rank sum test. Wilcoxon test z statistics and their accompanying p−values are listed in columns three and four. The null of an observation being plausibly drawn from either distribution cannot be rejected. This is true regardless of the number of OHIE eligible members in the household. Both tests suggest Oregon's randomization was successful not only from a baseline means perspective but across the conditional earnings distribution as well. Results suggest that any difference in the double eligible, dual survey responding households is likely related to attrition. However, given the lack of a clear conclusion from both diagnostic tests, the evidence against distributional equivalence is not strong. Furthermore, section A3.1 below demonstrates that restricting the sample to just the single eligible households holds no implications for the paper's results.

[3]Note there are only k−densities plotted for the single and double eligible households. There are some households with three eligible members, but the number is trivial.

A3. Additional Sensitivity Analyses

Outside of the extensions found in section five, one can consider the sensitivity of the paper's estimates to four additional robustness checks. This includes restricting the sample to just single eligible households, reducing the level of confidence to 90%, considering nonparametric alternatives to ITT QTE estimates, and the potential impact of other transfer programs with labor supply incentives (SNAP and TANF).

A3.1 Distributional Equivalence and the Influence of Double Eligible Households

In testing whether the identification assumption of distributional equivalence was satisfied in the previous section, the evidence for the double eligible households was inconclusive. To examine whether the double eligible households drive full sample QTE estimates, both the ITT and TOT regressions are re−estimated using only the single eligible households. Single eligible households clearly satisfied baseline distributional equivalence under either diagnostic test.

Figure IV plots the QTE estimates for the ITT estimates in Panel A and the TOT estimates in Panel B. For the sake of mapping quantiles to % FPL, Figure V displays a quantile plot for just the single eligible households. The magnitudes and qualitative implications are consistent with full sample results with three small exceptions. First, for the ITT results in Panel A, relatively more QTE point estimates differ from zero and the implied magnitude of these estimates is more negative than in Figure 2. Second, the range of statistically significantly positive QTE estimates between the $20^{\text{th}} - 30^{\text{th}}$ quantiles in Panel B is greater than in the full sample. Finally, additional upper tercile TOT QTE estimates are statistically significantly negative in Panel B compared to Figure 3. The pattern of consistently larger reductions in earnings also commences prior to the 100% FPL threshold for the single eligible Medicaid compliers. Overall, the results from the single eligible subsample regressions suggest that attrition concerns over the double eligibles are inconsequential to the paper's primary results.

A3.2 Full Sample and Single Eligible Only Regressions with 90% Confidence Intervals

As noted in the paper, some of the upper tercile TOT QTE estimates statistically insignificant at the 95% level of confidence are only marginally insignificant. Figures VI and VII reproduce results in Figures 2, 3, and IV using 90% confidence intervals. The main implications of choosing the 10% level for statistical inference are seen in Panel B of both Figures VI and VII. While many of the TOT QTE estimates immediately adjacent to the 100% FPL threshold and between the $80^{\text{th}} - 87^{\text{th}}$ quantiles were previously marginally insignificant at the 5% level, they are now statistically negative at the 10% level. This does not change the qualitative implications of the paper's results, outside of strengthening the empirical case for Medicaid representing a labor disincentive to relatively higher earning compliers.

A3.3 Nonparametric Alternatives to ITT Quantile Regressions

As non−parametric alternatives to the paper's ITT estimates, one can both visually examine the post−treatment empirical PDF's and test for distributional equivalence. Figure VIII plots the nonparametric empirical density functions of earnings for both treatment and control groups one year after lottery status notification. Panel A contains the single eligible k−densities while Panel B shows a similar PDF for the double eligibles. As the k−density graphs offer visual evidence of differences across the distributions, one should observe changes in probability mass from Figure II to Figure VIII consistent with static labor supply theory's predictions in Figure I. This includes both a shift left in probability mass for the treated group and bunching around the 100% federal poverty line (FPL) threshold.

Empirical density functions for both single and double eligible households are consistent with this prediction. For the single eligibles in Panel A, the probability mass between 175−200% FPL has fallen while more earners are located near 100% FPL. A similar pattern is observed with the double eligibles. Treated double eligible households are less likely to be found around 200% FPL and more likely to report earnings between 100−175% FPL. Relative to the baseline double eligible densities in Figure II, there are leftward shifts in the treated density. Control double eligibles were originally more likely to report earnings below 50% FPL than the lottery winners. After treatment, the densities below 50% FPL are equivalent. Finally, the lack of difference in the right tail for both household types is theoretically compelling. No effect at relatively high earnings levels coincides with Figure I's predictions for individuals like agent J.

While patterns of non−zero and zero treatment effects match theoretical priors, the visual evidence in Figure VIII is not definitive. The probability mass shifts for the treated group are small in magnitude for the single eligibles. While Medicaid's eligibility threshold affects earnings behavior above the federal poverty line, access to Medicaid does not visually change the treated density below 100% FPL for the single eligibles. As seen in columns 2 and 4 of Table A2, the K−S and Wilcoxon tests on the post−treatment earnings distributions also fail to reject distributional equivalence post−treatment at the 10% level. Nevertheless, as k−density analysis represents a nonparametric alternative to the paper's ITT estimates, it is unsurprising (and consistent with ITT results in Figure 2 of the paper) that the post−treatment distributions are statistically equivalent. It offers no evidence regarding Medicaid's distributional effect on OHIE compliers.

A3.4 Considering the Role of Alternative Transfer Payment Programs (SNAP, TANF)

It could be the case that access or take up of Medicaid through the OHIE lottery leads to an increased likelihood of participation in other public assistance programs such as TANF or SNAP. Finding that TANF or SNAP are responsible for observed labor supply reductions would not imply that Medicaid is irrelevant to earnings decisions. Rather, it would refine our understanding of the causal channel driving earnings decisions. Changes in labor supply can be related to Medicaid directly. Alternatively, the lottery could indirectly induce new Medicaid households to participate in other public assistance programs.

The plausibility of an indirect effect is supported by evidence from Baicker et al. (2014) and Chen, Flores, and Flores−Lagunes (2016). Both find that winning the OHIE lottery increased the likelihood of treated households participating in SNAP.[4] The presence of an indirect effect of Medicaid on labor supply holds two implications for interpreting this paper's estimates. First, following Chen et al., one could argue that the TOT results are the net effect of both Medicaid and receiving information on the safety net.[5] Alternatively, Baicker et al. suggest that any increased participation in other programs due to winning the lottery remains attributable to Medicaid participation and remains a direct effect of Medicaid.

In order to evaluate the influence of household participation in TANF or SNAP on my primary results, both the ITT and TOT results are re−estimated without SNAP or TANF participants. Clearly, the decision to participate in either SNAP or TANF is endogenous, but this test should be understood as considering how influential these observations are to the paper's primary results. If previously negative QTE estimates collapse to zero or appear otherwise sensitive to sample re−configuration, Medicaid's impact could be reappraised as primarily affecting labor supply indirectly.

[4] As noted in Baicker et al., lottery winners that registered for Medicaid in person were provided information on both SNAP and TANF.

[5] Ideally, in order to separate the effects of Medicaid and safety net information, Chen et al.'s bounds on the ATT in the presence of invalid instruments could be adapted to an IV quantile framework. Advancements in the partial identification literature may allow revisiting this concern in the future.

Plots of the QTE estimates from both series of regressions are found in Figure IX. There are two main observations to be drawn from the ITT results in Panel A and the TOT results in Panel B. First, the QTE estimates for both the ITT and TOT regressions below the 100% FPL threshold are consistently positive in sign and suggest relatively larger earnings increases than in the full sample. While the confidence intervals for these estimates are quite wide, a range of TOT QTE between the 10^{th} and 20^{th} quantiles are statistically positive. Second, while the TOT point estimates in Panel B do indicate earnings reductions for Medicaid participants in the top quintile of earnings distribution, these estimates are no longer statistically significant from zero.

Interpreting these results in favor of either direct or indirect effects remains unclear. Reversals of statistical significance alongside less negative QTE estimates are consistent with SNAP−related indirect effects. SNAP is more likely than TANF to have been the source of indirect effects for reasons of eligibility. Hoynes and Schanzenbach (2012) demonstrate that SNAP's implications for labor supply should only reinforce Medicaid's earnings disincentives. Removing SNAP participants as a result should lead to estimates revising upwards. Nonetheless, while this is true throughout much of the earnings distribution, it does not apply to TOT QTE's in the upper quartile of the earnings distribution. QTE point estimates in these regions still conform to static labor supply predictions, even though they are statistically insignificant.

References

[1] Baicker, Katherine, Amy Finkelstein, Jae Song, and Sarah Taubman. "The Impact of Medicaid on Labor Market Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment." *American Economic Review: Papers and Proceedings*. 104, no. 5 (2014): 322−328.

[2] Chen, Xuan, C. Flores, and Alfonso Flores−Lagunes. "Bounds on Average Treatment Effects with an Invalid Instrument, with an application to the Oregon Health Insurance Experiment." Unpublished Manuscript (2017). Accessed at http://conference.iza.org/conferencefiles/IFAU2016/chenx9567.pdf.

[3] Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker. "The Oregon Health Insurance Experiment: Evidence from the First Year." *Quarterly Journal of Economics*. 127, no. 3 (2012): 1057−1106.

[4] Hoynes, Hilary Williamson, and Diane Whitmore Schanzenbach. "Work incentives and the food stamp program." Journal of Public Economics 96, no. 1 (2012): 151−162.

[5] Imbens, Guido W., and Jeffrey M. Wooldridge. "Recent Developments in the Econometrics of Program Evaluation." Journal of Economic Literature 47, no. 1 (2009): 5−86.

Appendix Tables and Figures

Figure I: Medicaid–Augmented Labor Supply Model

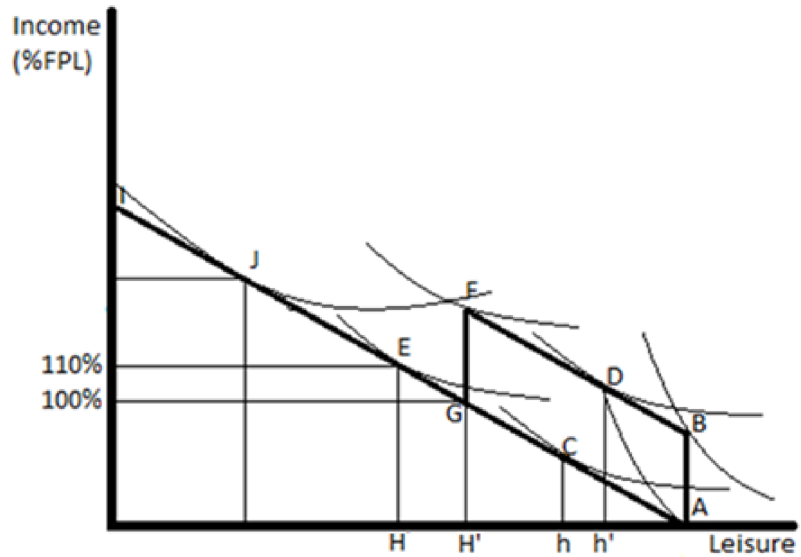
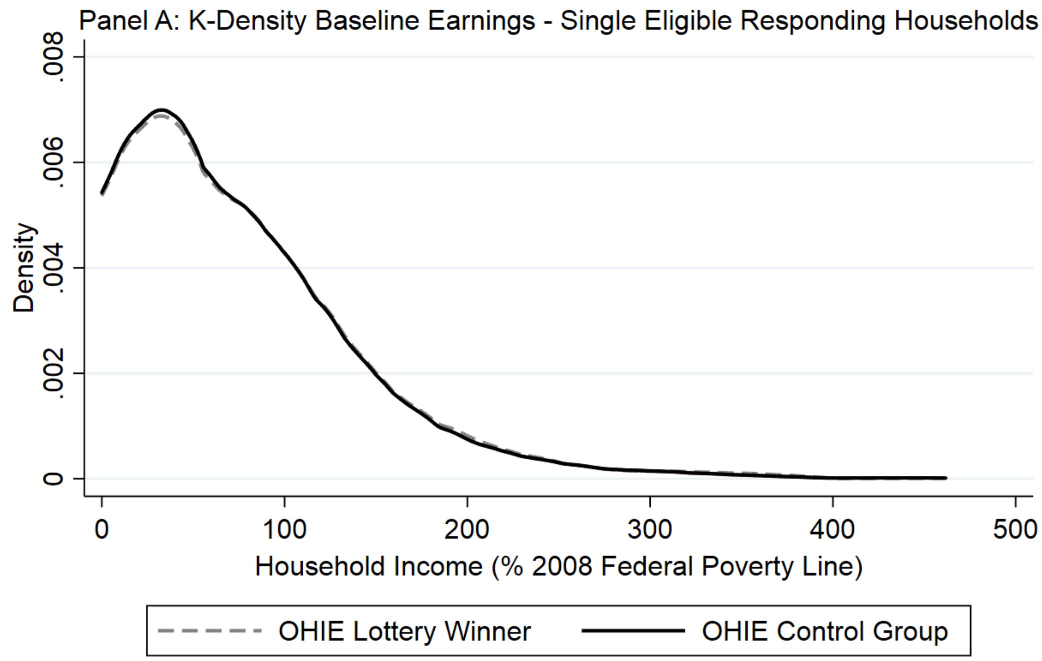
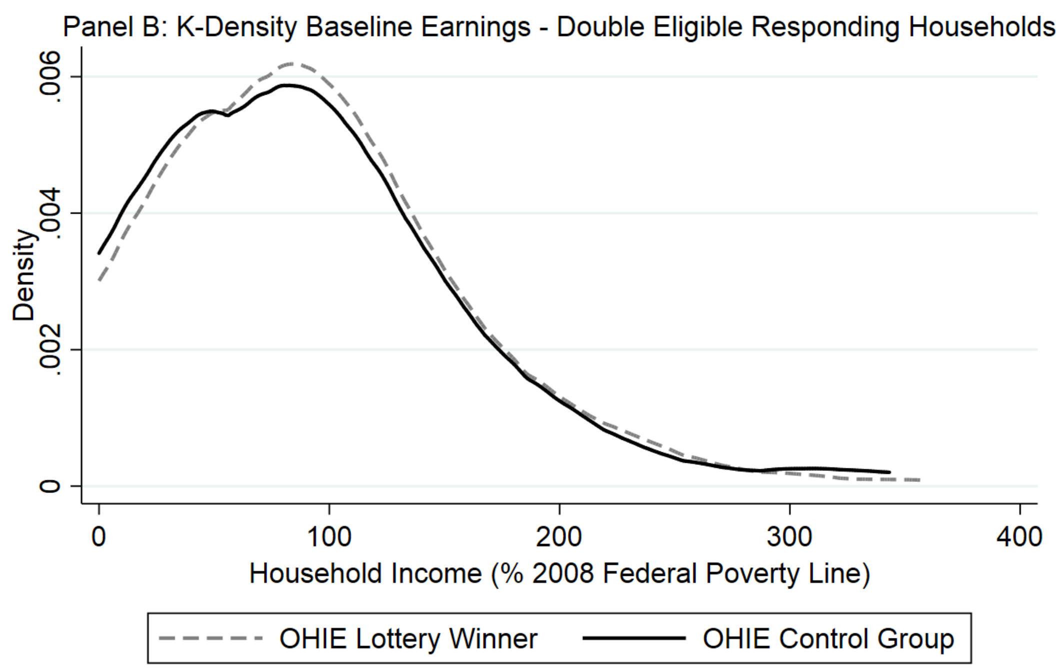


Figure I plots the predicted effect of Medicaid on earnings, local to different regions of the low-income budget set. Similar to the OHP-Standard Medicaid program in Oregon, the cutoff for eligibility is at 100% FPL in the figure.

Figure II: Baseline Empirical Earnings Distributions by Lottery Status and Number of OHIE Eligible Household Members



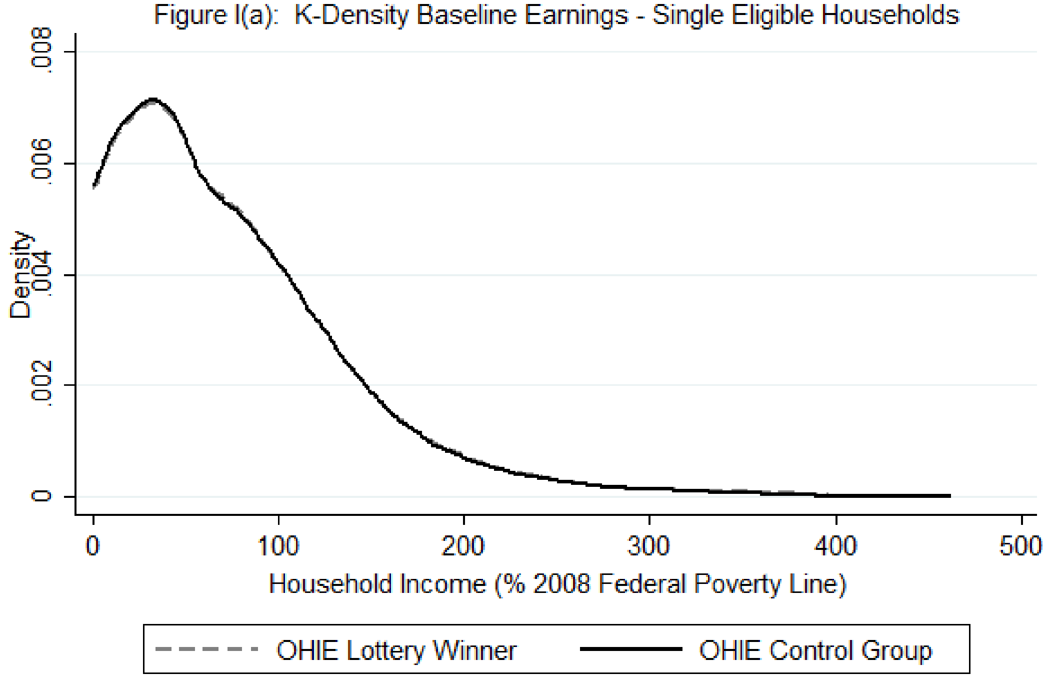
Note: Plot displays kernel density functions for treatment and control households with one OHIE eligible member that respond to the baseline and 12 month surveys. Kernel density functions are estimated using a 25 unit bandwidth specification.



Note: Plot displays kernel density functions for treatment and control households with two OHIE eligible members that respond to the baseline and 12 month surveys. Kernel density functions are estimated using a 25 unit bandwidth specification.

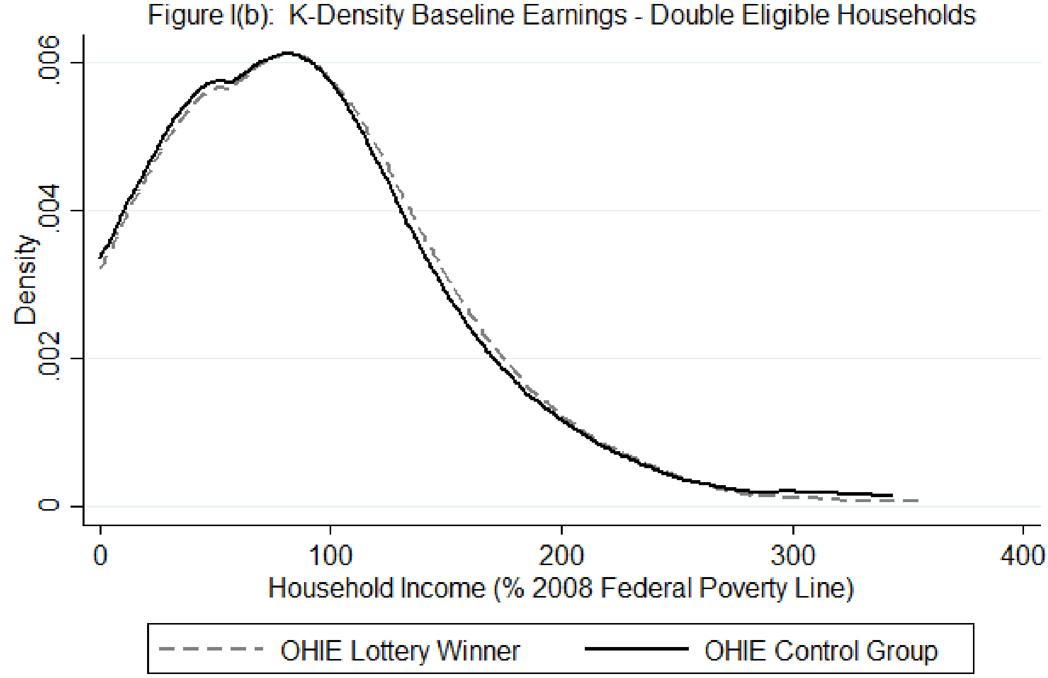
Figure III: K-Density Baseline Earnings for Baseline Survey Responding Households

Figure I(a): K-Density Baseline Earnings - Single Eligible Households



Note: Plot displays kernel density functions for treatment and control households with one OHIE eligible member to the baseline survey. Kernel density functions are estimated using a 25 unit bandwidth specification.

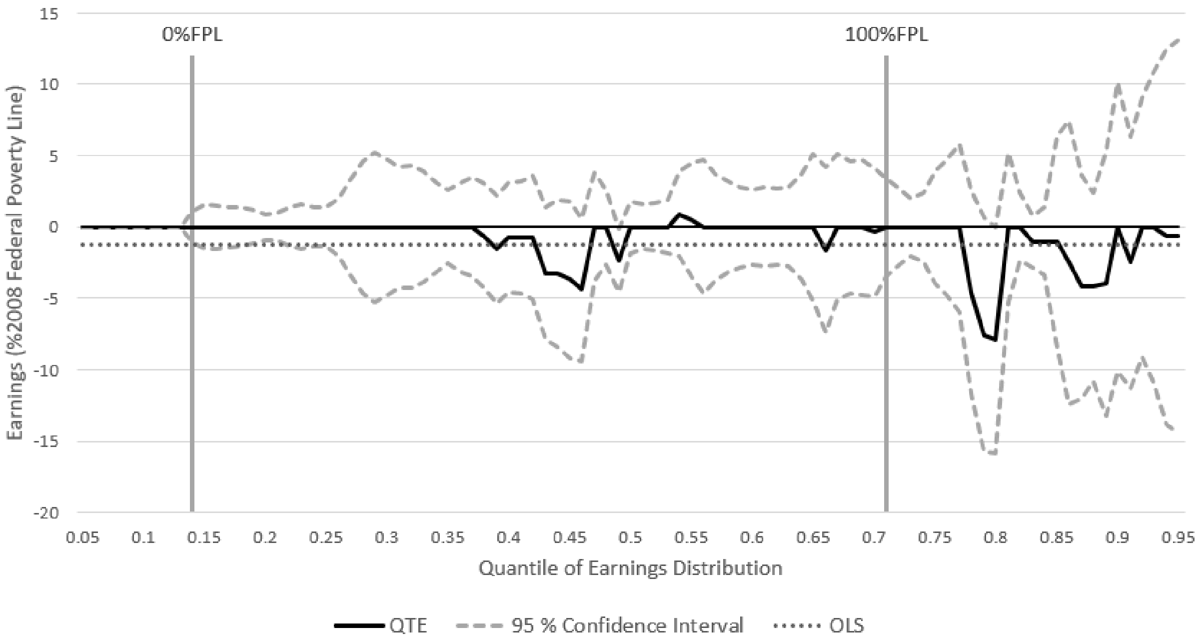
Figure I(b): K-Density Baseline Earnings - Double Eligible Households



Note: Plot displays kernel density functions for treatment and control households with two OHIE eligible members to the baseline survey. Kernel density functions are estimated using a 25 unit bandwidth specification.

Figure IV: One Year Quantile Treatment Effects for Single Eligible Households

Panel A: One Year Intent-to-Treat (ITT) Quantile Treatment Effects of OHIE Medicaid Lottery Assignment on Earnings for Single Eligible Earning Households



Notes: Quantile treatment effects of winning the OHIE lottery for households with only one OHIE eligible member are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 95% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the ITT model. The OLS estimate is obtained from the same ITT specification. It is not statistically significant based on standard errors clustered at the household level. (N = 14,859, Treated = 6,902, Control = 7,957).

Panel B: One Year Treatment-on-the-Treated (TOT) Quantile Treatment Effects of Medicaid Participation on Earnings for Single Eligible Earning Households

Notes: IV Quantile treatment effects of single eligible household participation in OHP Standard through the OHIE lottery are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 95% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the TOT model. The 2SLS LATE estimate is obtained from the same TOT specification. It is not statistically significant based on standard errors clustered at the household level. (N = 14,859, Treated = 6,902, Control = 7,957).

Figure V: Quantile Plot of 2008 Earnings Distribution for Single Eligible OHIE Households

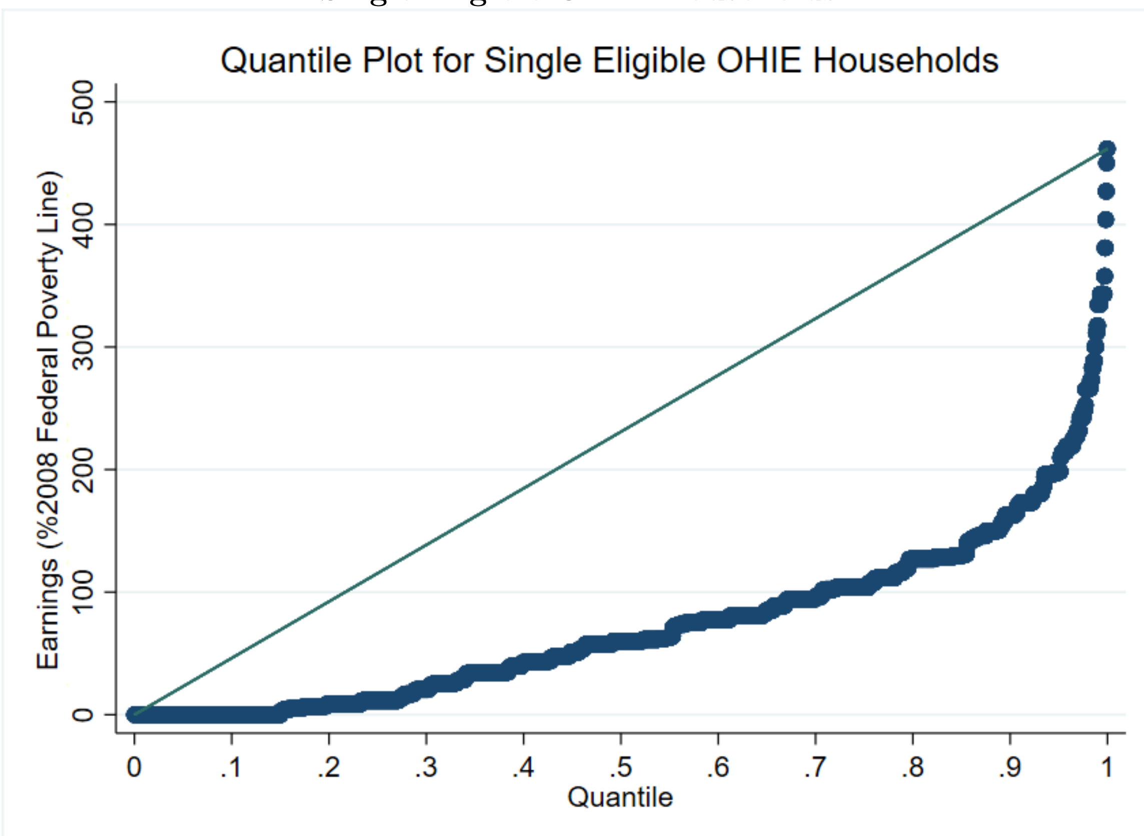
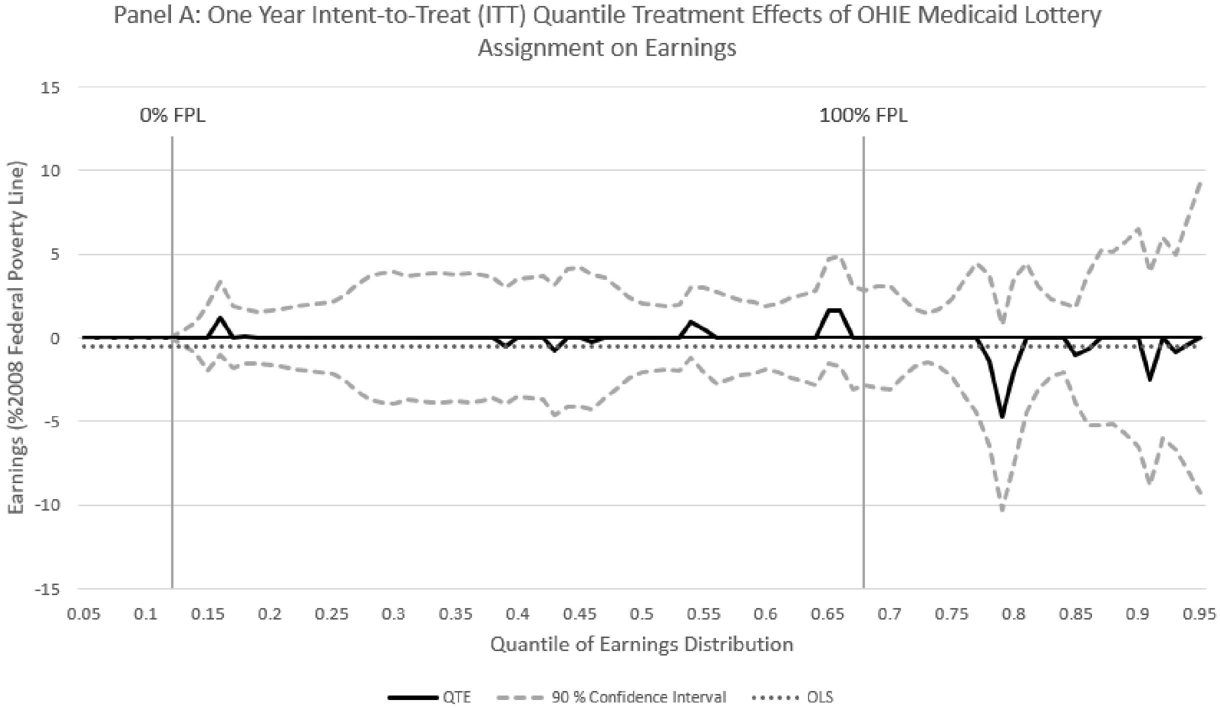
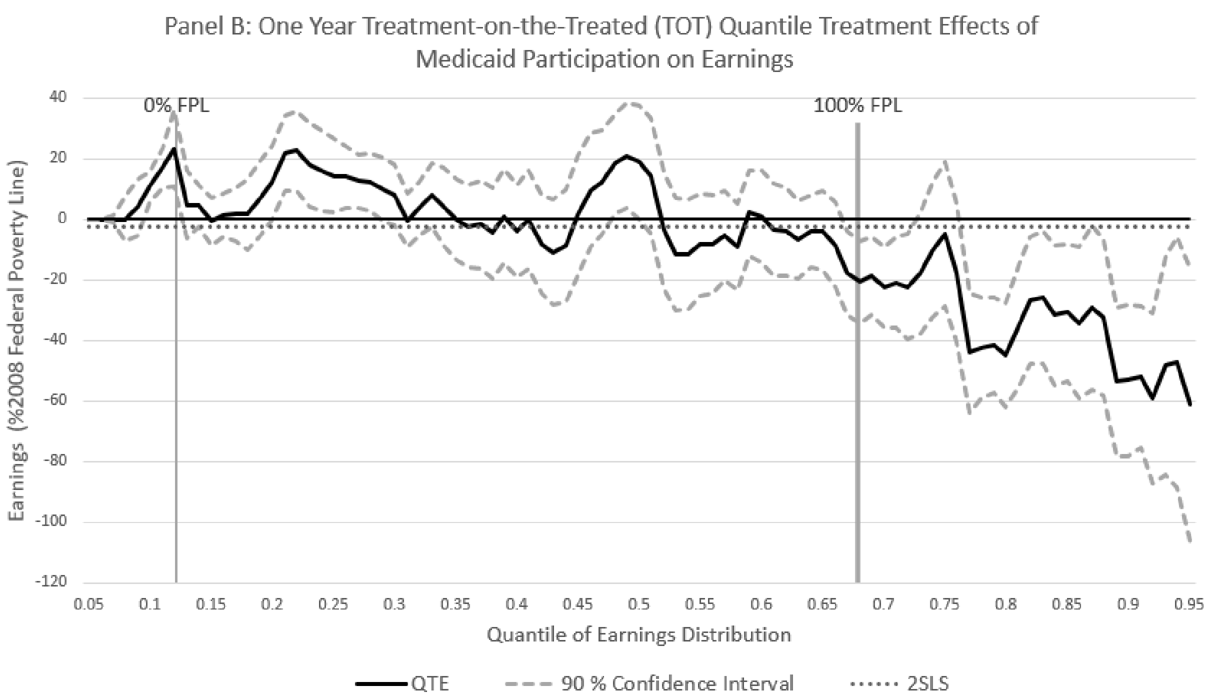


Figure V maps household earnings levels with only one OHIE eligible household member, measured as % of the 2008 Federal Poverty Line, to respective quantiles in the earnings distribution.

Figure VI: One Year Quantile Treatment Effects with 90% Confidence Intervals



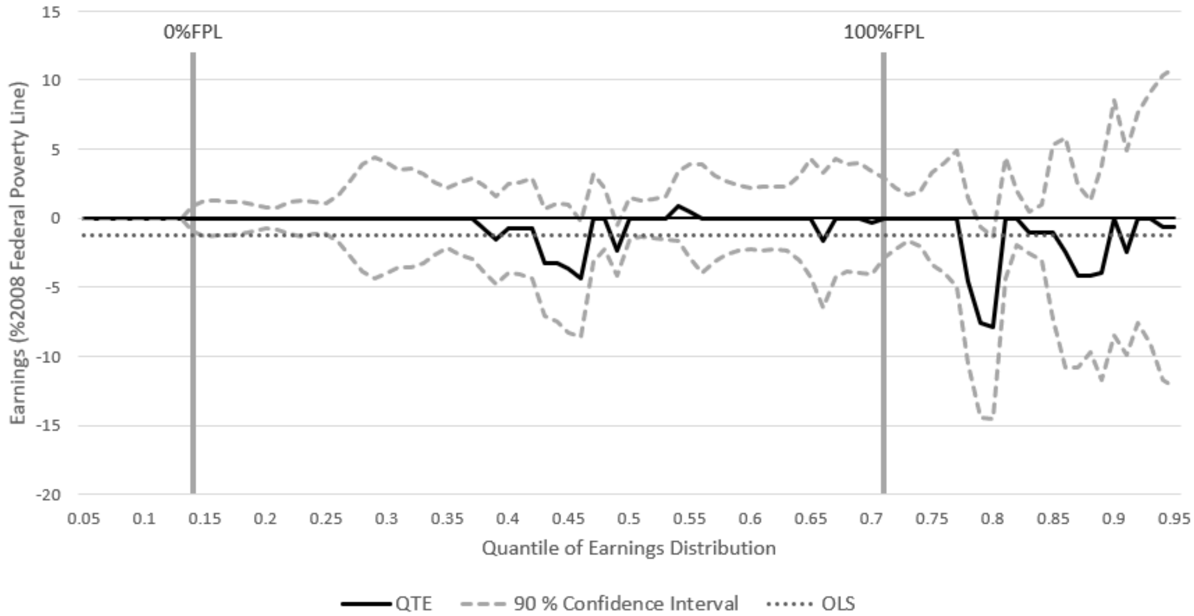
Notes: Quantile treatment effects of winning the OHIE lottery are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 90% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the ITT model. The OLS estimate is obtained from the same ITT specification. It is not statistically significant based on standard errors clustered at the household level. (N = 16,566, Treated = 8,147, Control = 8,419).



Notes: IV Quantile treatment effects of participation in OHP Standard through the OHIE lottery are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 90% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the TOT model. The 2SLS LATE estimate is obtained from the same TOT specification. It is not statistically significant based on standard errors clustered at the household level. (N = 16,566, Treated = 8,147, Control = 8,419).

Figure VII: One Year Quantile Treatment Effects for Single Eligible Households with 90% Confidence Intervals

Panel A: One Year Intent-to-Treat (ITT) Quantile Treatment Effects of OHIE Medicaid Lottery Assignment on Earnings for Single-Eligible Earning Households



Notes: Quantile treatment effects of winning the OHIE lottery for households with only one OHIE eligible member are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 90% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the ITT model. The OLS estimate is obtained from the same ITT specification. It is not statistically significant based on standard errors clustered at the household level. (N = 14,859, Treated = 6,902, Control = 7,957).

Panel B: One Year Treatment-on-the-Treated (TOT) Quantile Treatment Effects of Medicaid Participation on Earnings for Single-Eligible Earning Households

Notes: IV Quantile treatment effects of single eligible household participation in OHP Standard through the OHIE lottery are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 95% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the TOT model. The 2SLS LATE estimate is obtained from the same TOT specification. It is not statistically significant based on standard errors clustered at the household level. (N = 14,859, Treated = 6,902, Control = 7,957).

Figure VIII: Post-Treatment Empirical Earnings Distributions by Lottery Status and Number of OHIE Eligible Household Members

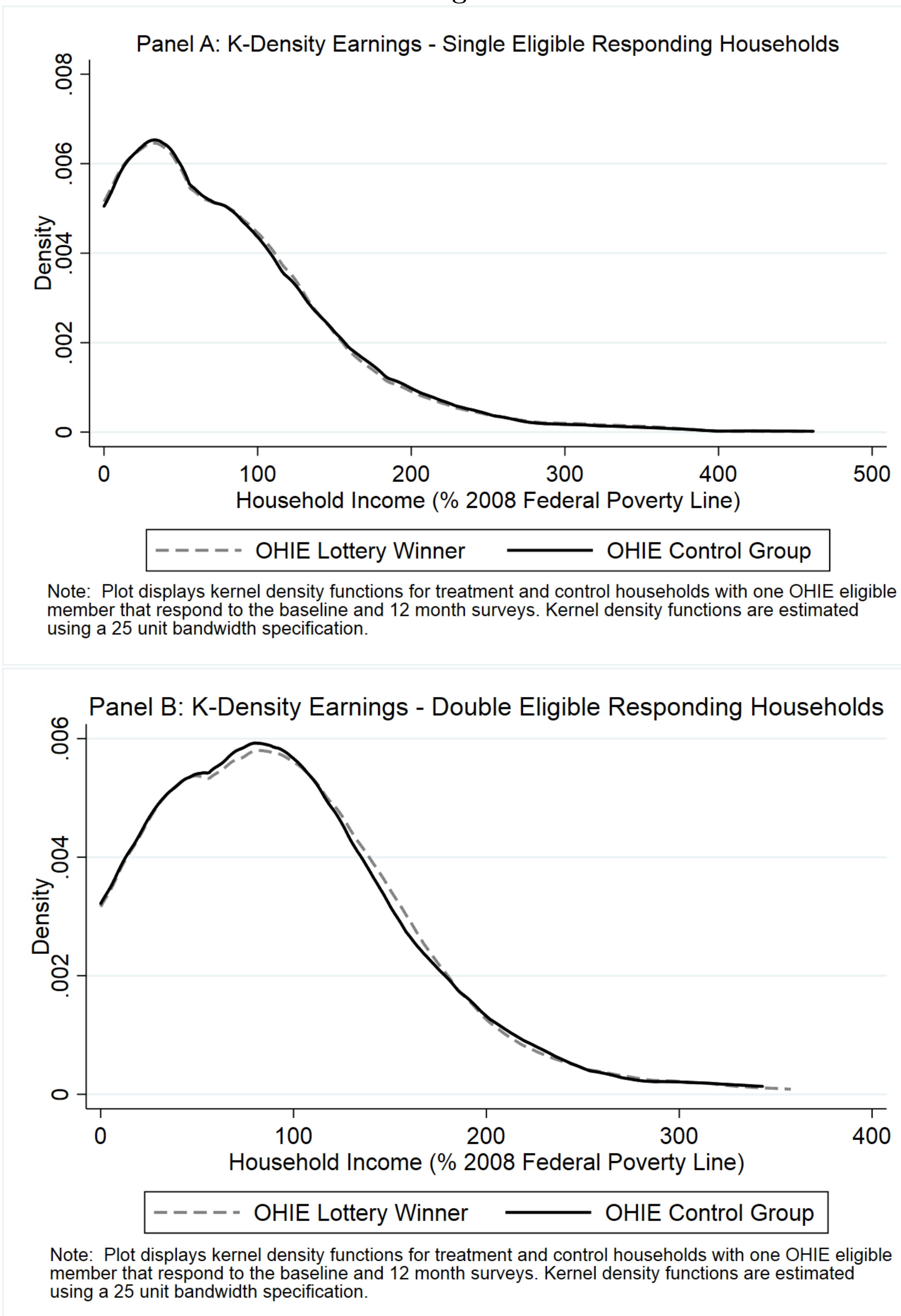
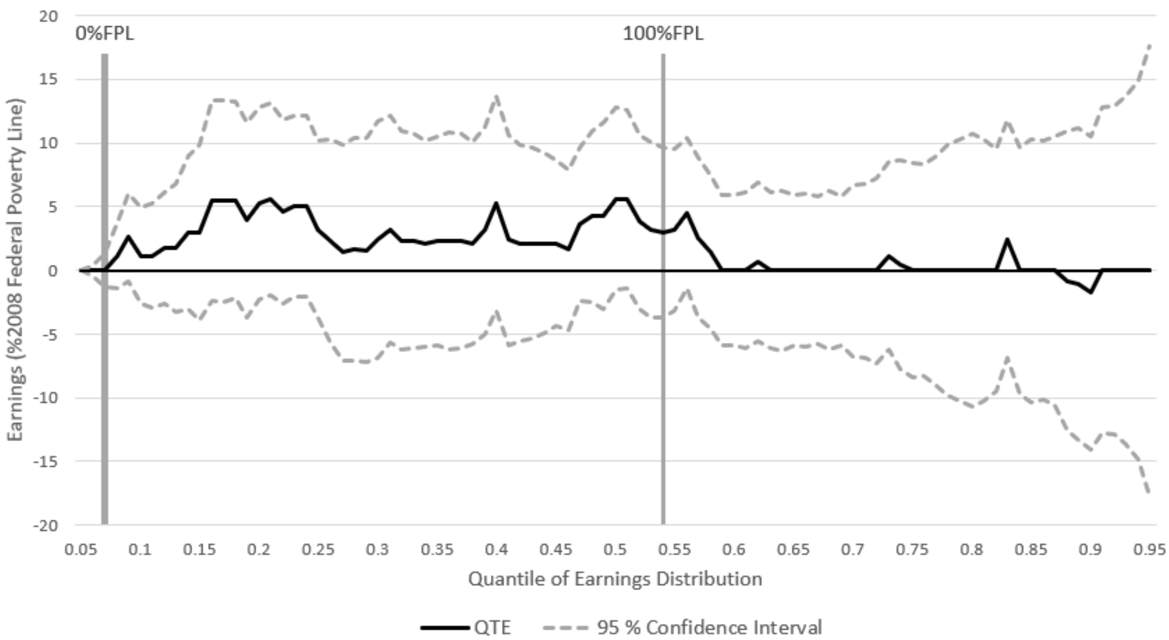


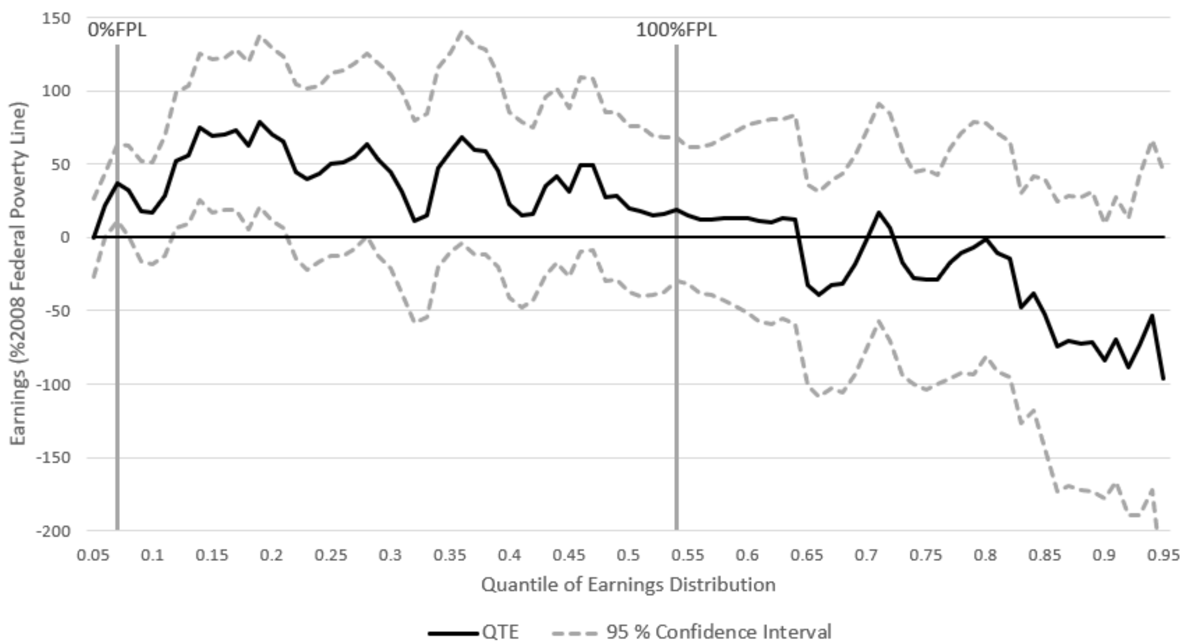
Figure IX: One Year Quantile Treatment Effects Excluding Households Participating in SNAP/TANF

Panel A: ITT QTE Estimates for All OHIE HouseholdsNot Participating in SNAP or TANF



Notes: Quantile treatment effects of winning the OHIE lottery for households are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Earnings are scaled by household size into %FPL measures. Estimation is conducted using a subset of OHIE participants that did not simultaneously indicate participation in SNAP or TANF. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 95% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation of the ITT model. (N = 5,512, Treated = 2,646, Control = 2,866).

Panel B: TOT QTE Estimates for All OHIE Households Not Participating in SNAP or TANF



Notes: IV Quantile treatment effects of participation in OHP Standard through the OHIE lottery are estimated from the 5th to 95th quantiles of the conditional earnings distribution. Estimation is conducted using a subset of OHIE participants that did not simultaneously indicate participation in SNAP or TANF. Earnings are scaled by household size into %FPL measures. Survey responses are scaled using sampling weights from Finkelstein et al. (2012). 95% confidence bands are nonparametrically calculated from 500 bootstrap replications for each quantile's estimation. (N = 5,512, Treated = 2,646, Control = 2,866).

Table A1: Testing Distributional Equivalence of Baseline Earnings by Lottery Status

	K-S D Statistic (1)	P-Value (2)	Wilcoxon z-Statistic (3)	P-Value (4)
Panel A: Baseline and 12 Month Survey Respondents by Household Size				
One Eligible Member (N=9,893, C=5,348 T=4,545)	0.0126	0.830	-0.599	0.5492
Two Eligible Members (N=4,416, C=1,896 T=2,520)	0.0390*	0.075	-1.633	0.1024
Three Eligible Members (N=32, C=5 T=27)	0.4815	0.282	-0.784	0.4332
Panel B: Baseline Survey Respondents by Household Size				
One Eligible Member (N=15,414, C=8,163 T=7,251)	0.0125	0.586	-0.546	0.5852
Two Eligible Members (N=6,966, C=2,956 T=4,010)	0.0195	0.540	-0.978	0.3281
Three Eligible Members (N=7, C=32 T=39)	0.2455	0.879	0.202	0.8399

Notes: Column one reports the D statistic from the Kolmogorov-Smirnov (K-S) Test. Column three reports the z-statistic from the Wilcoxon-Mann-Whitney rank sum test. Statistical significance of each diagnostic test is indicated by stars and p-values. Statistical significance is indicated as follows: *** = $P < 0.01$; ** = $P < 0.05$; * = $P < 0.10$.

Table A2: Testing Distributional Equivalence of Post-Treatment Earnings by Lottery Status

	K-S D Statistic	P-Value	Wilcoxon z-Statistic	P-Value
	(1)	(2)	(3)	(4)
Baseline and 12 Month Survey Respondents by Household Size				
One Eligible Member	0.0138	0.702	0.220	0.8258
Two Eligible Members	0.0317	0.222	-0.437	0.6617
Three Eligible Members	0.2315	0.992	-0.177	0.8594

Notes: Column one reports the D statistic from the Kolmogorov-Smirnov (K-S) Test. Column three reports the z-statistic from the Wilcoxon-Mann-Whitney rank sum test. Statistical significance of each diagnostic test is indicated by stars and p-values. Statistical significance is indicated as follows: *** = $P < 0.01$; ** = $P < 0.05$;* = $P < 0.10$.