WIC Participation and Relative Quality of Household Food Purchases: Evidence from FoodAPS

by

Di Fang*

Michael R. Thomsen

Rodolfo M. Nayga, Jr.

Aaron M. Novotny

April 2018

* Di Fang is Assistant Professor, Michael R. Thomsen is Professor, and Rodolfo M. Nayga Jr. is Distinguished Professor and Tyson Chair in Food Policy Economics all in the Department of Agricultural Economics and Agribusiness, University of Arkansas. Aaron M. Novotny is a PhD candidate in the Department of Economics, University of Arkansas.

This research was supported by USDA grant no. 59-5000-5-0115 to the National Bureau of Economic Research, entitled, "Using FoodAPS for Research in Diet, Health, Nutrition, and Food Security." The views expressed are those of the authors and not necessarily those of the Economic Research Service, Food and Nutrition Service, or the U.S. Department of Agriculture.

WIC Participation and Relative Quality of Household Food Purchases: Evidence from FoodAPS

Abstract

This paper examines the effect of the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) on the quality of household food purchases using the National Household Food Acquisition and Purchase Survey (FoodAPS) and propensity score matching. Purchase quality was measured using a healthy purchasing index (HPI). Findings indicate that WIC foods explain the improvement in quality of food purchases, not self-selection of more nutrition-conscious households into the program. Households participating in WIC have a higher HPI in comparison to eligible non-participating households. Importantly, this difference is driven entirely by WIC participating households who redeemed WIC foods during the interview week. There was no significant difference between WIC-participants who did not redeem WIC foods and eligible non-participants. The paper also examines whether geographic barriers limit WIC participation. Locations of WIC clinics were added to the already detailed FoodAPS information on food store locations. There is no evidence in this sample that access to clinics is adversely affecting participation nor is there evidence that HPI depends on supermarket access. Finally, a supervised machine learning process supports our main conclusion that WIC-provided foods are the driver of increased HPI, not self-selection of healthier households into the program.

Introduction

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is the nation's third largest food assistance program (Morgan 2015) but focuses narrowly on pregnant, postpartum, or breastfeeding women; infants; and children up to five years of age. The program provides food assistance, nutrition education, breast feeding support, and referrals to health and other services (USDA Food and Nutrition Service (FNS) 2016). WIC foods are intended to be supplemental and address nutritional gaps in the recipients' diets.[1] Participation in the program involves frequent visits to a WIC clinic for other services. Typically, at least one clinic appointment is required within a three-month period.

The program began in 1974 and its reach has grown with time. Recent estimates indicate that WIC serves half of US infants and close to 30 percent of children, pregnant women, and postpartum women (Oliveira and Frazão 2015). WIC targets lower income households, those with incomes below 185 percent of the federal poverty guidelines, but participants in some higher-income households are adjunctively eligible by having previously qualified for participation in another assistance program such as Medicaid (Thorn et al. 2015). That said, in 2014 only a small fraction, less than two percent, of participants were from households with income over 185 percent of the federal poverty level. The overwhelming majority, nearly three quarters, were in households with income below 100 percent of the poverty threshold (Thorn et al. 2015).

That WIC is associated with improvements in birth, health, and nutritional outcomes has been documented in numerous studies, with birth outcomes receiving the most attention.

[1] A recent estimate is that the average dollar value of WIC foods is $45 per recipient per month (Tiehen and Frazão 2016)

Reviews of this literature include Owen and Owen (1997), Fox et al. (2004), Devaney (2010) and Black et al. (2012). That said, selection bias is a concern in this work because nearly all studies compare participants to eligible non-participants (Gordon and Nelson 1995; Besharov and Germanis 2001). If participation is more attractive to those who are concerned about nutrition and/or otherwise exhibit behaviors contributing to better health, then the beneficial effects of the program could be overstated or spurious. Conversely, there is evidence that participants in WIC are more likely to have characteristics associated with poor health outcomes relative to eligible nonparticipants, which could result in published findings that understate beneficial effects of the program (Bitler and Currie 2005a).

Nevertheless, the emerging evidence is that the effects of WIC on birth outcomes are robust to selection bias (Currie and Rosin-Slater 2015). Recent studies have examined birth outcomes using empirical strategies designed to address selection bias and continue to show that WIC is beneficial. Strategies include models with maternal fixed effects (Sonchak 2016; Currie and Rajani 2014) and instrumental variables models (Gai and Feng 2012). Others exploit the staggered deployment of WIC across counties during the early stages of the program (Hoynes, Page and Stevens 2011) and compare outcomes from mothers transitioning into and out of the program over multiple births (Figlio, Hamersma, and Roth 2009).

There has been less direct attention on the role of selection bias in estimating the effect of the program's nutritional outcomes. Overwhelmingly, studies that have examined nutritional outcomes find that WIC participation is associated with improvements in dietary quality, nutrient intakes, and/or biochemical indicators of nutritional adequacy across a number of different datasets and time periods (see Fox et al. 2004; Black et al. 2012). Because of the selection issue, it is possible to argue that those attracted into the program would have purchased the supplemental WIC foods anyway, and that despite the large body of evidence showing a strong association between the program and improved diet, the program's actual benefits in terms of meaningfully augmenting nutrition are limited.

Assessing how selection might explain the nutritional outcomes of WIC is especially important now. WIC foods account for about 70 percent of program costs (Oliveira and Frazão 2015) and WIC is one of several nutrition and health programs targeted for cuts in a recent White House budget proposal (Aisch and Parlapiano 2017). The supplemental nature of WIC foods has led some to question whether the WIC food packages are sufficiently meaningful to the overall diet of pregnant women to alter birth outcomes and even whether the birth outcomes being considered would be sensitive to the level of nutrition supplementation that WIC provides (Joyce, Gibson, and Coleman 2005). Others have conceded the possibility that the food packages may be an incentive that induces pregnant women to participate in the program and that it is other program features, such as education and health referrals, that could be responsible for mitigating the likelihood of poor outcomes (Bitler and Currie 2005b). Thus, evidence on whether WIC foods matter is important to understanding the mechanisms by which the program leads to better health.

The question we address in this paper is whether participation in WIC meaningfully alters food choices in a way that would be conducive to improvements in diet. We address this question using USDA's National Household Food Acquisition and Purchase Survey (FoodAPS). FoodAPS provides a relatively small but nationally representative sample of US households and includes information about where households shop for food, the availability and types of food stores in the communities where these households are located, and household eligibility and participation for food assistance programs (USDA Economic Research Service (ERS) 2017a). Additionally, FoodAPS contains detailed information about factors that may influence food purchases such as socioeconomic characteristics of the household, number of individuals residing within the household, and racial composition. An advantage of the FoodAPS data is that they permit us to look at food purchases directly and see how they differ among participants who did and did not use WIC to purchase foods. Since the sample was not stratified by time of month or date of delivery of WIC benefits, the samples of WIC households using or not using their WIC benefits should be nearly random.

To determine how WIC contributes to the nutritional quality of household food purchases, we use an adaption of the Healthy Eating Index (HEI), a broadly accepted measure for the overall quality of an individual's diet and apply it to each household's food purchases. Because we are assessing purchases and not dietary intake, we term this measure the healthy purchasing index (HPI). FoodAPS provides a detailed record for participating households during an interview period (typically a week). These purchase records include information about what foods were purchased, where they were purchased, and whether purchases were made using benefits from nutrition assistance programs such as WIC. Moreover, FoodAPS includes nutrient values for the food items contained in the household's purchase record. This permits application of the HEI over these purchases to obtain the HPI.

Given the lack of data on a valid instrumental variable that could be used to help identify the effect of WIC participation, we use propensity scores to match WIC participating households to eligible non-participating households. We then estimate the average treatment effect on the treated (ATT) and show that WIC participation is associated with modest but statistically significant improvements in nutritional quality of household food purchases. This finding is consistent with earlier work showing an association between WIC participation and improved nutritional outcomes. Next, we compute the ATT among participants who redeemed WIC benefits for foods during the interview week and those who did not. We find that the improvement in nutritional quality of food purchases is driven entirely by households who redeemed WIC foods. We found no meaningful program effect on the purchases of WIC participating households who did not redeem WIC foods. In sum, we present evidence that WIC foods are the most plausible explanation for earlier findings of a positive association between WIC participation and nutritional outcomes. Moreover, we find no evidence to support the contention that this association can be explained by systematic differences among WIC and comparable non-WIC households.

We conduct several robustness checks on this finding. First, food retailers who accept WIC benefits as payment carry healthier foods. It is plausible that differences in shopping venue could explain this finding. Hence, we redo the analysis using a sample comprised only of households who shopped at a WIC approved retailer during the interview week and reach the same conclusion. WIC foods continued to explain the difference in purchase quality in this follow-up analysis.

Second, our finding could be due to a secondary selection issue wherein some WIC-participating households have characteristics that make them more likely to fully redeem WIC benefits. To assess this, we exclude food items procured on shopping trips where WIC accounted for a majority of the value of a household’s purchases, recompute the HPI, and redo the analysis. When the WIC shopping events are excluded, the ATT among those households who redeemed WIC foods is no longer significant. Again, this supports the conclusion that WIC foods explain the improvement in nutritional quality, not self-selection of more nutrition-conscious households into the program.

While propensity score matching is a well-established method, one weakness is that the researcher must specify a relationship between the likelihood of program participation and observed covariates. As a third robustness check, we use an "honest" random forest that places no restrictions on model complexity but that does penalize overfitting to derive the ATTs (Athey, Tibshirani, and Wager 2017; Athey 2017; Wager and Athey 2017). This generalized random forest approach uses machine learning methods to identify proximity of participating and non-participating households and thereby sidesteps the requirement of a researcher-supplied functional specification.

A secondary aim of this paper is to use the rich information FoodAPS contains about the commercial food environment to understand whether geographic barriers impact WIC participation. Specifically, we look at whether inadequate access to WIC clinics limits participation and whether the HPI differs meaningfully for participants without access to supermarkets. We find little evidence from the FoodAPS data that these barriers pose significant hurdles to program participation and effectiveness.

Data and Methods

FoodAPS contains a total of 4,826 households who completed the survey between April 2012 and January 2013 (USDA ERS 2017a). Among the 4,826 households who took part in FoodAPS, there were 1,007 households with at least one member who was categorically eligible for WIC and who met other program requirements for income or adjunctive eligibility through participation in Medicaid or other qualifying assistance program. Our focus is on this subsample of FoodAPS households. Of these 1,007 eligible households, 461 households were participants in WIC. Households recorded purchases in food-at-home and food-away-from-home food diaries. We further restricted the WIC-eligible sample to those households with at least one food-at-home event during the interview period and those that constituted complete cases over all measures reported in Table 1. As shown in Table 1, our analysis sample includes 928 households. Of these, 505 households were eligible for WIC but did not participate in the program and 423 participated. Of the 423 participating households, 152 used WIC benefits on one or more purchase occasions during the interview period and 271 did not.

FoodAPS contains information about household characteristics that may affect WIC participation and food choice. As shown in Table 1, these characteristics include educational attainment, monthly income, marital status, presence of different categories of WIC-eligible individuals, and household racial and ethnic composition. Most of these measures are based on the characteristics of individuals in the household, which are then aggregated up to the household level.[2]

In addition, we are able to examine two geographical barriers: access to supermarkets and access to WIC clinics. Supermarket access could be one barrier that influences both participation in WIC and food choices in general. To designate households without easy access

[2] Educational attainment reflects the highest attainment of anyone in the household. The number of WIC infants and WIC children is based on individuals aged 0 to 1 and 1 to 4, respectively. A binary measure is used to indicate the presence of a WIC eligible woman. This is set to one if there was a pregnant woman in the household, there was an infant being breastfed, or if there was a birth within the household within the last three months. There is the potential for underreporting WIC-eligible women since WIC provides postpartum benefits for six months but FoodAPS flags birth events within the past three months.

to supermarkets, we use USDA's tract-level measure indicating limited access to supermarkets based on vehicle travel (USDA ERS 2017b). Second, as demonstrated by Rossin-Slater (2013), access to WIC clinics is another determinant of WIC take-up. Since the FoodAPS dataset does not provide information on clinic access, we assembled a list of WIC clinic locations. This involved collection of data on clinic locations across numerous state WIC agencies. To correspond temporally to the FoodAPS data collection period, we used 2012-2013 locations if available, but only the 2015-2016 locations were available for some states. We supplied the resulting geocodes for WIC clinics to USDA-ERS personnel who spatially joined the clinic locations to the FoodAPS households and provided a file with radial distances from each household to the nearest WIC clinic within that household's state of residence. We then measured clinic access as a binary variable taking the value of one if the household was within one (ten) miles of a clinic and located in an urban (rural) tract.[3]

Measuring the Healthy Purchasing Index (HPI)

As noted earlier, the HPI measure we use for nutritional quality is based on the HEI. Specifically, our measure is based on the HEI-2010, which reflects diet quality in terms of conformance to the 2010 Dietary Guidelines for Americans (USDA FNS 2010). The HPI measure we use differs primarily in that (a) it is computed over food purchases as opposed to food intake and (b) it is measured at the household as opposed to the individual level. Nevertheless, this measure still provides a baseline measure of a household's ability to meet dietary guidelines from its food purchases. The HEI, and by extension the HPI used here, assesses 12 dietary components (Guenther et al. 2013). These include nine adequacy components

[3] Our use of one and ten mile radii for urban and rural tracts is analogous to the definition used to identify limited access to supermarkets for purposes of defining food deserts (see USDA ERS 2017b).

and three moderation components. The adequacy components reflect (1) total fruit, (2) whole fruit, (3) total vegetables, (4) greens and beans, (5) whole grains, (6) dairy, (7) total protein foods, (8) seafood and plant proteins, and (9) fatty acids. The moderation components include (1) refined grains, (2) sodium, and (3) empty calories. A higher value for each component of the index indicates a healthier nutrient intake. The overall HPI measure computed over these 12 categories ranges from 0 to 100. Again, higher values of the overall index indicate healthier food purchases.

The FoodAPS survey collected detailed information about all foods purchased by the household over the course of seven days. The primary respondent in each household participated in two in-person interviews and up to three telephone interviews (Ver Ploeg et al. 2015). The dataset contains information on food items purchased or otherwise acquired, including brand, and package size, which allowed items to be matched to the USDA Food and Nutrient Database for Dietary Studies or the USDA National Nutrient Database for Standard Reference (USDA ERS 2017a). Economists at the USDA's Economic Research Service developed computer code and intermediate datasets that aggregated these nutritional values into the HEI components and computed an overall index value at the household level. These datasets and computer code were then made available for our use in this research and were used to compute the HPI index we use as the outcome measure in this study.

Characteristics of the WIC and Eligible Non-WIC Samples

As shown in Table 1, the mean HPI score for WIC and eligible non-WIC households are virtually identical at 50.259 and 50.388, respectively. The last two columns of the table break the WIC households down into those who redeemed WIC benefits during the sample period and those who did not. Here the difference in HPI is meaningful. The average HPI computed over those who redeemed WIC foods is much higher at 55.958 in comparison to the average of 47.062 computed over those who did not. This is not surprising because foods provided by WIC help participants meet dietary guidelines. Nevertheless, the apparent importance of WIC foods to the magnitude of the index is striking and will receive further attention below.

One thing that is noteworthy from Table 1 is that there are important differences between the WIC and eligible non-WIC samples. The average number of WIC-eligible individuals is similar between the two groups but on average, WIC households contain higher numbers of infants and eligible women than did non-WIC households. WIC households have lower levels of educational attainment, a lower percentage of married couples, and much lower household incomes. The proportion of Hispanic households is much higher among the WIC sample and the proportion of African American households is slightly higher. Finally, access to WIC clinics is higher among WIC households (40.4 percent compared to 35.8 percent) but a larger proportion of WIC households did not have access to a supermarket (22 percent compared to 14.5 percent).

To assess households' subjective evaluation of dietary quality, FoodAPS included the primary respondent's self-assessment of whether the household is following a healthy diet. As shown in Table 1, a slightly higher percentage, 40.9 percent, of WIC households reported a healthy diet in comparison to 38.2 percent of the non-WIC households. Nevertheless, given the differences between the WIC and non-WIC samples, the interesting question remains whether WIC truly improves the healthiness of food purchase for participants or whether participants self-select into WIC because of preferences for healthier foods such as those provided by WIC. To answer this question, we use a matching algorithm to estimate the ATT of WIC participation.

Matching WIC Households to Eligible Non-WIC Households

As noted earlier, estimating the impact of WIC on nutritional quality is difficult with observational data such as FoodAPS because the treatment selection (in this case, WIC participation) is often influenced by subject characteristics. Consequently, baseline characteristics of treated subjects could differ systematically from those of untreated subjects. Therefore, to understand the effect of participation in WIC on HPI, we must first account for the systematic differences in baseline characteristics between WIC participants and eligible non-participants.

Matching methods provide a way to reduce selection bias among observational data (Rosenbaum 2002; DiPrete and Gangl 2004). The goal is to find a group of non-treated individuals who are similar to the treated individuals in all baseline characteristics – then focus attention on estimating the effect of interest and consider all variables other than the treatment variable as potentially confounding. Balancing the vector of characteristics across treatment reduces the influence of confounding variables. Therefore, matching mimics a randomized experiment (conditional on a set of observables) so that the effect of the treatment is established (Drichoutis, Nayga, Lazaridis 2009).

Matching methods are discussed at length in Rosenbaum and Rubin (1983). The matching algorithm we use constructs an artificial control group among eligible non-WIC households that have similar characteristics as those of WIC participants. Let T_i indicate the treatment, which equals one if household *i* participated in WIC (treated case) and zero if household *i* is WIC-eligible but does not participate in WIC (control case). Define HPI outcomes as Y_{0i} and Y_{1i} for the associated treatment status 0 and 1. The treatment effect for an household i can be written as: $t_i = Y_{1i} - Y_{0i}$. However, we do not know t_i because we can only observe the outcome of either $t_1$or t_0 (we only observe $Y_i = t_i Y_{1i} + (1 - t_i)Y_{0i}$), but not both. Therefore, we can estimate the average treatment effect on the treated (ATT) as $t_{ATT} = E(Y_1|T = 1) - E(Y_0|T = 1)$. Understanding the effect of WIC participation on those who ultimately participated is the relevant policy question in our study, not the effect of WIC participation averaged across all households regardless of whether or not they participated in WIC. Hence, we focus on the estimation of ATT in our analyses.

Notice that the term $E(Y_0|T = 1)$ is not observed because we do not observe the WIC effect of households who are not on WIC. Moreover, if one tries to substitute this with $E(Y_0|T = 0)$, it would lead to self-selection bias (Drichoutis, Nayga, Lazaridis 2009). We can assume that selection into treatment depends on observable covariates as long as the following two strong ignorability assumptions in treatment assignment are satisfied: (1) $(Y_{1i}, Y_{0i}) \perp T_i|X_i$; and (2) $0 < p(X) < 1$. The first condition implies that selection is solely based on observable characteristics and that all variables influencing treatment assignment and potential outcomes simultaneously are observed by the researcher (Caliendo and Kopeinig 2008). The second ensures a common support (to rule out perfect predictability of treatment given ***X***) between the treatment and control groups. Rosenbaum and Rubin (1983) further demonstrated that under the assumptions of strong ignorability, treatment and control groups are exchangeable. The average treatment effect for the treated is estimated as $t_{ATT} = E\{E(Y_1|X, T = 1) - E(Y_0|X, T = 0)|T = 1\}$, where the outer expectation is taken over the distribution of baseline covariates in the treated group (Rosenbaum and Rubin 1983). Therefore, outcome analysis on the matched data tends to produce unbiased estimates of treatment effects due to reduced selection bias through the balancing of the distributions of observed covariates.

In this study, we use nearest neighbor propensity score matching (PSM). Propensity score matching has been popularly applied in economics, statistics and medical research (Hong and Yu 2008; Ye and Kaskutas 2009; Wyse, Keesler, and Schneider 2008; Staff et al. 2008). PSM forms matched sets of treated and untreated subjects who share a similar probability to be treated (Rosenbaum and Rubin 1983; 1985). We use a logistic regression model to estimate the propensity score. Specifically, WIC participation status is regressed on observed household characteristics. Afterwards, we use the “Matching” package in R to obtain the matched samples (Sekhon 2011). The algorithm we use matches each treatment household to a control household with replacement. We restricted matches to nearest neighbor within a caliper as small as 0.01, but our central findings and the overall quality of matches are robust to caliper restrictions. Consequently, the matched samples we report below include all WIC participating households in the sample or subsample that had common support. Once the matched sample is formed, the treatment effect can be estimated by directly comparing outcomes between treated and untreated subjects in the matched sample.

Results

In this section, we first present results from the logistic regressions that we use to model WIC participation. Second, we assess the quality of our matches by checking the balance between the WIC and eligible non-WIC samples. Third, we then present the ATT of WIC participation and assess Rosenbaum sensitivity of the ATT to hidden bias or unobserved heterogeneity. Fourth, we assess robustness of our main findings by repeating the analysis for a sample containing only households who shopped at a WIC-approved retailer and again using a modified HPI that excludes the majority of WIC purchases. We then present information on differences in ATT by self-reported healthiness of purchases and by access to supermarkets. Finally, we present the ATT estimates derived through an alternative generalized random forest algorithm.

Determinants of WIC Participation

Marginal effects from the logit models used to form the matched samples are reported in Table 2. The model in Table 2 includes state fixed effects (not reported). These fixed effects are included because earlier studies show that states with stricter WIC eligibility rules have lower take-up (Bitler, Currie, and Scholz 2003; Swann 2010). To formally test whether the existence of state effects matters, we conducted a log-likelihood ratio test comparing models with and without the state effects and reject the null hypothesis of no state effects.

Table 2 shows that households with higher income and the highest levels of educational attainment are less likely to participate in WIC. Hispanic households are more likely to participate in WIC. Household composition is also important. Specifically, households with eligible infants and women are more likely to participate in WIC. These findings are largely consistent with earlier work that has examined WIC participation. For example, there is evidence that WIC take-up is lower for children age 1 to 4 (Bitler, Currie, and Scholz 2003; Tiehen and Jacknowitz 2008) and higher among eligible postpartum women (Tiehen and Jacknowitz 2010). Earlier work also shows higher rates of participation among Hispanic households (Bitler and Currie 2004; Bitler, Currie, and Scholz 2003). WIC take-up is higher among socially disadvantaged women (Tiehen and Jacknowitz 2010; Swann 2010; Bitler, Currie, and Scholz 2003).

We find no evidence that access to WIC clinics affects participation among households in the FoodAPS sample. This is in contrast to Rossin-Slater (2013) who finds that access to clinics increases WIC food benefit take-up. However, Rossin-Slater (2013) has a much larger sample from a single state and exploited information about opening and closing of clinics, information which we were unable to obtain for the more geographically diverse FoodAPS sample. That we do not find a significant effect of clinic access could be due to differences in program delivery and clinic access across the many states represented in FoodAPS. The model in Table 2 does not include a covariate measuring supermarket access, but we look further at the issue of supermarket access in a follow-up analysis below.[4] In sum, we do not find evidence that geographic barriers (i.e., in terms of clinic access) meaningfully affect WIC participation in this sample.

Assessing the Quality of Matched Samples

As explained above, the goal of propensity score matching is to obtain a dataset that is similar to one that would result from a randomized experiment. For this reason, we want the distribution of covariates to be the same between the matched treated and control groups. One way to check this is assess the balance post-match. We use the standardized difference measure proposed by Rubin (1991). For each explanatory variable in the logit model, the standard difference of the sample means in the treated and matched controls are presented in Figure 1. The overall mean difference before matching lies between 0.4 percent and 80.4 percent for WIC households and eligible non-WIC households. The bias is reduced to between 0.4 percent and 13.0 percent in the post-match sample.

[4] We did not include supermarket access to avoid the causal loop. The majority of supermarkets take WIC. It is unclear whether households take up WIC because they have access to supermarkets or whether they shop at a supermarket that takes WIC because of WIC participation.

Since the covariates include not only continuous variables, but also binary variables, a Kolmogorov-Smirnov test based on 2,000 bootstrap iterations is employed to provide correct coverage as recommend by Sekhon (2011). Based on the Kolmogorov-Smirnov test, before matching, the unbalanced variables are household income, households of Hispanic ethnicity, the existence of a WIC eligible woman in the household, and the number of WIC-eligible children in the household. After matching, these differences are reduced and there is no longer significant imbalance between WIC households and eligible non-WIC households. We also use a two-sample t-test, as proposed by Caliendo and Kopeinig (2008), to check if there are significant differences in covariate means between treated and matched. These two-sample t-tests provide additional evidence that covariate balance is achieved at the 5 percent level (see Appendix).

Finally, we test the assumption of common support by checking the distribution of the propensity scores for the treatment and control groups, as exhibited in Figure 2. As shown in Figure 2, almost all the treated observations could be matched with non-zero propensity score control observations and we restricted the matched samples to the region with common support.

Effect of WIC on Nutritional Quality of Household Purchases

Table 3 presents the ATT estimates for the effect of WIC participation regardless of whether or not households redeemed WIC vouchers during the interview week, the effect of WIC participation for households who had a WIC food redemption, and the effect of WIC participation for households who did not have a WIC food redemption. The control group in all these analyses comprises those who are eligible but non-WIC participants. We estimate a positive and statistically significant ATT of 2.742 for WIC participation. As noted above, the HPI ranges from 0 to 100. Given that the average HPI value for WIC participants is about

50.259, this estimate suggests that WIC participation improves the nutritional value of purchases by about 5.5 percent. This finding is consistent with earlier work showing that WIC participation is associated with improvements in diet quality, nutrient intakes, and biochemical indicators of nutritional adequacy (e.g., see review by Fox et al. 2004).

To shed light on the importance of WIC foods, Table 3 also reports ATT estimates for the sample of WIC participating households who used WIC benefits to redeem foods during the interview week and those households who did not. The ATT is much higher at 9.443 among the households who redeemed WIC foods. In contrast, the ATT for the sample of WIC households who did not redeem WIC foods is -0.843.

It is not surprising that the ATT is higher when WIC foods are redeemed. After all, foods eligible for purchase through WIC are those that help recipients meet dietary guidelines. The important finding is that there is no evidence that WIC participation improves nutritional quality among those participating households who do not redeem benefits. If households with healthier food preferences self-selected into the program, we should see a higher ATT even when WIC benefits are not redeemed. The ATT from this group is small and not statistically different from zero.

As noted above, the HPI we use to assess quality of household purchases is an aggregate of the 12 HEI components. Table 4, reports ATT estimates for each of these components estimated from the matched sample of WIC households redeeming WIC benefits. Table 4 suggests that households redeeming WIC foods scored significantly higher on total fruit, whole fruit, whole grains, dairy, seafood and plant protein, refined grains, and empty calories. This is not surprising because these categories are emphasized/deemphasized in the WIC food packages.

Earlier evidence shows reduced intakes of fats and added sugars among WIC participants (Basiotis, Kramer-LeBlanc, and Kennedy 1998; Wilde, McNamara, and Ranney 1999; Kranz and Siega-Riz 2002; Siega-Riz et al. 2004). The 2009 changes to the WIC food packages resulted in increased purchases of whole grain products (Oh, Jensen, and Rahkovsky 2016).

Rosenbaum Bounds to Assess Hidden Bias

When referring to hidden bias, we assume that some characteristics are unobserved and are not in the vector of covariates used in the matching model. Propensity-score matching estimators are based on the assumption that selection is on observable characteristics. This means that conditional on the observed covariates, the process by which units are selected into treatment is independent of unmeasured variables that affect the outcome variable. In order to estimate the extent to which such selection on unobservable or hidden bias may affect the estimates, we conducted a Rosenbaum bounds sensitivity analysis (Rosenbaum 2002; DiPrete and Gangl 2004; Drichoutis, Nayga, Lazaridis 2009). This method assesses the sensitivity of the significance levels of the ATT and estimates the magnitude of hidden bias it would take to change inference assessments from statistical significance to insignificance. Details about computing Rosenbaum bounds can be found in Rosenbaum (2002) and DiPrete and Gangl (2004). We used the "rbounds" package in R to conduct the sensitivity analysis (Keele 2010). Tables 3 and 4 present Rosenbaum's gamma, the measure of hidden bias that that could potentially switch an inference decision at the 5 and 10 percent critical values.

Gamma is interpreted as the magnitude by which an unobserved variable would need to affect the odds ratio of treatment in order to cause an inference decision to switch from being significant to insignificant. As the Rosenbaum test reveals, our ATT from the sample of all WIC participants switches from being statistically significant to insignificant at a gamma value of 1.13 and 1.17 at the 5 and 10 percent levels, respectively. This indicates that the estimate would remain significant at the 10 percent level in the presence of hidden bias up to 17 percent. Table 3 shows that the significance of the large ATT estimated for the sample of WIC households who redeemed WIC benefits is very robust to hidden bias with gamma values of 2.82 and 3.05 at the 5 and 10 percent levels, respectively.

The statistically significant ATTs on the component measures estimated from this subsample in Table 4 are also robust to hidden bias. The refined grain component is the most sensitive with a gamma of 1.22 corresponding to the 5 percent critical value. Gammas for the other component scores range from 1.28 to 2.25 at this critical value. Hence, the ATTs on most of the significant component measures would remain significant at the 5 percent level even in the presence of substantial hidden bias.

Additional Evidence on the Importance of WIC Foods

To summarize, the evidence presented so far supports the conclusion that the improvements in nutritional quality attributable to WIC participation are driven by WIC food packages and are not simply a reflection of selection bias. This conclusion is strengthened by the fact that the FoodAPS sample was not stratified by time of month or date of delivery of WIC benefits. For this reason, the samples of WIC households redeeming and not redeeming their WIC benefits should be nearly random. Indeed, these two groups of households have very similar characteristics as exhibited in the last two columns of Table 1.[5]

[5] The two groups of households differ statistically only on the household members classified as African American (at 0.05 level).

However, there are two potential issues that deserve further attention. First, WIC foods must be redeemed at WIC approved retailers. If these retailers stock healthier foods in general, then differences in shopping venue could account for some of the improvements in nutritional quality attributable to the subsample that redeemed WIC foods. Second, our finding could reflect a secondary selection problem wherein some households who enroll in the program are systematically more likely to only partially redeem food benefits. This could occur if shopping venues available to the household stock some but not all of the foods in the WIC package or if some households deem some WIC foods to be undesirable.

To address the first issue, we restricted the sample to include only participating and non-participating households who shopped at a WIC-approved retailer during the interview period. This restriction resulted in the removal of three households from the eligible but non-participating sample and four households who were on WIC but did not redeem WIC benefits during the interview week. Thus, it is unlikely that venue explains the results because the overwhelming majority of our sample shopped at a WIC approved retailer. However, the exclusion of these few households did alter the matched samples as can be seen by comparing the numbers of observations in Table 5 to those in Table 3. Nevertheless, as shown in Table 5, The ATT estimates are very close to those reported above. For WIC participants without regard for redemptions, the ATT estimate was 2.671. For those who redeemed and did not redeem foods during the interview week, the estimates were 9.385 and -0.963, respectively. Thus, in comparison to Table 3, no materially different conclusions are reached from the analysis summarized in Table 5.

To address the second issue, we re-estimated the ATT from the original matched samples in Table 3 but using an HPI that excludes items from shopping events where WIC redemptions accounted for more than 50 percent of the total expenditures. In FoodAPS, each shopping event is flagged as to whether WIC benefits were redeemed during the purchase event and the dollar value of WIC redemptions is indicated[6]. Unfortunately, FoodAPS does not provide item-level information on which items were purchased with WIC benefits and which were purchased using other forms of payment. However, of the 273 WIC purchase events, the overwhelming majority were solely WIC events. These could be identified by the fact that the total value spent on the shopping occasion was equal to the dollar value of WIC redemptions.[7] In only 28 shopping event cases did WIC redemptions account for less than half of the total value of the food purchases. Thus, by excluding majority-WIC shopping events, we effectively remove most WIC foods from the HPI calculation.

As shown in Table 6, when this revised HPI is used as the outcome, there is no longer a significant WIC effect. Among all WIC participants without regard for redemptions, the ATT estimate is -0.087. For those who redeemed and did not redeem foods, the ATT estimates are 1.420 and -0.843, respectively. The estimate for households who did not redeem WIC foods matches that in Table 3 because these households acquired no foods through WIC during the interview week and so the HPI score remains unchanged for these households. The magnitude of each estimate is close to one index point and not statistically different from zero, suggesting that

[6] In October 2009, the USDA revised the WIC food package and introduced cash-value vouchers (CVV) for fruits and vegetables. Since FoodAPS does not separately identify CVV, the relatively small value recorded for fruits and vegetables expenditures when WIC benefits were used may reflect misreporting of the vegetables and fruits expenditures when using the CVV (National Academies of Sciences, Engineering, and Medicine. 2017). Therefore, our estimates of the effect of WIC benefits may be underestimated (i.e., a lower bound).

[7] The number of shopping events will not match the number of households because households can have multiple shopping events during the interview period.

there is no difference between WIC and eligible non-WIC households once WIC foods are effectively removed from the measure of nutritional quality. Again, this reinforces the conclusion above that WIC foods are the best explanation for observed improvements in dietary outcomes associated with the program, not systematic differences in the characteristics or behaviors of participating and eligible non-participating households.

Additional Insights on WIC from FoodAPS

FoodAPS includes the response to a self-assessed question about whether the household is following a healthy diet. As reported in Table 1, the proportion responding yes to this question is similar across WIC participants and eligible non-participants but is a bit lower among those who redeemed WIC benefits during the reporting period in relation to those who did not. As a follow-up, we matched those with yes and no responses to this question separately to eligible nonparticipants to obtain ATT estimates for each group. For those responding yes, the ATT is 3.720 (Std. Err=1.710) and is statistically significant at the 5 percent level. For those responding no, the ATT estimate is about 1.5 index points lower at 2.137 (Std. Err=1.474) and is not significantly different from zero. Nevertheless, this is not a substantial difference in the ATT estimates and the fact that these estimates are similar provides some additional context to the selection issue we explore above.

Another concern is whether households in lower income neighborhoods without access to nearby supermarkets may benefit less from nutritional programs like WIC. Given the geographic component of FoodAPS, we explore the heterogeneity that may exist because some households have limited access to supermarkets. We match 93 WIC households with limited supermarket access and 330 WIC households with supermarket access to eligible non-WIC households. We estimate the ATTs for each subgroup. The ATT estimate from the limited access subgroup is 2.340 (Std. Err=1.971) and is not statistically different from zero while the ATT from the subgroup without limited access is 2.984 (Std. Err=1.417) and is significant at the 5 percent level. Given the similarity of these ATT estimates, there is no compelling evidence from this sample that nutritional improvements from WIC are adversely affected by supermarket access.

Robustness Check Using Machine Learning

With propensity matching, we assumed a certain function between the likelihood of WIC participation and the covariates. In reality, this relationship can be complex and unknown. To further check the robustness of our findings presented above, we apply an alternative method, a supervised machine learning approach called the "honest" random forest that places no restrictions on model complexity (but penalizes over-fitting) to derive the ATTs (Athey, Tibshirani, and Wager 2017; Athey 2015; Wager and Athey 2017). An "honest" tree splits a randomly selected subsample for the use of model structure estimation. Therefore, the asymptotic properties of treatment effect estimates within the splits are the same as if the partition had been exogenously given (Athey, Tibshirani, and Wager 2017; Athey and Imbens 2016; Wager and Athey 2017). We use the "causal_forest" command developed by Athey, Tibshirani, and Wager (2017), which is available in the R package "grf". Since we are not concerned about the interpretability of the prediction model, we included additional covariates to check if they can improve estimation.[8] Essentially, these covariates help us detect similarities among WIC participants and among eligible non-participants and assign them to different groups

[8] These additional covariates included alternative measures of supermarket accessibility, household use of nutrition information, number of meals consumed at home during the interview period, and additional indicators of household economic status in addition to those covariates listed in Table 2. The causal trees were also estimated using only the covariates in Table 2, but results (not reported) are similar to those reported below in Table 7.

(leaves). This method relies on a type of residual-on-residual regression in the leaves to eliminate the effect of confounding. Intuitively, if our main findings from the propensity score matching analysis are a result of the true data-generating process, we should not see different effects when we use an alternative functional form or approach provided through this machine-learning based method.

Our causal tree estimates are reported in Table 7 and are consistent with the main results discussed above. Specifically, we find similar ATTs on the overall WIC sample (ATT=2.267), the sample that redeemed WIC (ATT=7.720), and the sample that did not redeem WIC (ATT=-0.903). These results confirm our conclusion that only those who redeemed WIC observe a significant WIC effect, or that the WIC package is the driver of increased HPI. Our examination of the HPI that excluded majority WIC purchase events yields an ATT of -0.109, which is close to zero, statistically insignificant, and similar to that reported in Table 6. This supports our conclusion that WIC-provided foods are driving the improvement in diet and that this improvement is not simply a reflection of healthier households selecting into the program. When reviewing the effect of WIC on the HPI components, the list of components with positive and statistically significant improvements is similar the that reported above and includes total fruit, whole grain, dairy, seafood and plant protein, sodium, and empty calories. The total vegetables component was negative in Table 4 and is negative and significant in Table 7. The estimate for the total-vegetable component is notable because at the time the FoodAPS data were collected, white potatoes were not permitted for WIC purchases (Oliveira and Frazão 2015). White potatoes are also the most widely consumed fresh vegetable in the United States (USDA-ERS 2017c). The exclusion of white potatoes from WIC-eligible fresh vegetables is a plausible explanation for the negative ATT on the total vegetable component.

Conclusions

Children of low-income households tend to lag behind other children on a wide range of health outcomes. They also tend to be food insecure and have inadequate intake of important nutrients. As previously discussed, the role of WIC in improving birth, health and nutritional outcomes has been studied extensively. The findings of these studies generally suggest that WIC participation is associated with improved outcomes. However, most earlier work showing beneficial effects of WIC on diet have been unable to convincingly determine whether the observed beneficial effects on diet are due to the program or to self-selection into the program.

We addressed this important topic using the FoodAPS data and found that households participating in WIC have higher HPI value in comparison to eligible non-participating households. While this finding is consistent with earlier findings linking WIC to improvements in diet quality, we found that this difference is driven entirely by households who redeemed WIC foods during the interview week. Importantly, we did not find any difference between WIC participating households that did not redeem WIC foods and WIC-eligible but non-participating households. Moreover, there was no difference in purchase quality when majority WIC shopping events were excluded from the HPI calculation. These findings suggest that WIC foods explain the improvement in relative quality of household food purchases, not self-selection of more nutrition-conscious households into the program. This is the key contribution of the present study.

To be clear, it is not surprising to find that household purchases were healthier when WIC program benefits were used to purchase foods. After all, foods eligible for purchase under WIC are specifically designed to help beneficiaries meet dietary guidelines. Instead, the evidence against self-selection being an explanation for improvements in the quality of purchases lies in two main findings. First, WIC households had healthier purchases only when benefits were redeemed. During weeks when benefits were not redeemed, the nutritional quality of purchases among WIC households was no different from that of eligible non-participating households. Second, among those redeeming WIC benefits, the improvement in purchase quality disappeared once WIC purchase events were excluded from the purchase quality index. After exclusion, there was no longer a difference between those households who redeemed WIC foods and eligible nonparticipating households. Taken together, the evidence presented here suggests that foods provided by WIC can explain most if not all of the improvement in the quality of food purchases. There is little if anything left that can be explained by systematic differences in the characteristics of participating and non-participating households.

Access to the FoodAPS data permitted us to look at food purchases, the first link in a chain connecting WIC participation to better child and maternal health outcomes. Earlier studies have looked at links further along this chain such as diet quality, indicators of nutritional adequacy, and birth outcomes. The focus on purchases is important because we were able to examine the point in the chain where behaviors related to self-selection into the program would be easiest to detect. Had there been evidence of healthier purchases either among participating households who did not redeem benefits or among the non-WIC purchases of households who did redeem benefits, it would have supported an argument that WIC households would have been likely to make healthier purchases regardless of the program. We found evidence of neither. Thus, our findings suggest that WIC-provided foods are central to the beneficial program outcomes documented in earlier studies. Moreover, our findings on purchase quality align nicely with earlier findings on diet quality.

Overall, the findings of this study point to the importance of the WIC program in helping participants acquire the foods needed for a healthier diet. The ability of WIC to continue serving eligible households could be curtailed, however, if current proposals to significantly cut WIC funding push through. This issue becomes even more relevant for eligible households when considering that WIC is different from other welfare programs in that it is not an entitlement program. Not everyone who is eligible for WIC assistance is guaranteed to receive it because the reach of the program depends on adequate funding.

References

Aisch G., and A. Parlapiano. 2017. How Trump's budget would affect every part of government. *New York Times* May 23. Available https://www.nytimes.com/interactive/2017/05/23/us/politics/trump-budget-details.html (accessed November 9, 2017).

Athey, S. and G. W. Imbens. 2016. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences* 113 (27):7353-7360.

Athey S., J. Tibshirani,, and S. Wager. 2017. Generalized random forests. arXiv:1610.01271v3 [stat.ME] (accessed February 12, 2018).

Athey, S. 2017. The impact of machine learning on economics. In Agrawal, A.K., J. Gans, and A. Goldfarb, Eds. *Economics of Artificial Intelligence*. University of Chicago Press (forthcoming). Available http://papers.nber.org/books/agra-1 (accessed February 12, 2018).

Basiotis, P.P., C.S. Kramer-LeBlanc, and E.T. Kennedy. 1998. Maintaining nutrition security and diet quality: The role of the food stamp program and WIC. *Family Economics and Nutrition Review* 11 (1/2):4-16.

Besharov, D.J., and P. Germanis. 2001. *Rethinking WIC: An Evaluation of the Women, Infants, and Children Program*. Washington, DC: AEI Press.

Bitler, M. P., and J. Currie. 2004. *Medicaid at Birth, WIC Take-Up, and Children's Outcomes.* WR -172. Santa Monica, CA: RAND Corporation. Available https://www.rand.org/pubs/working_papers/WR172.html (accessed February 12, 2018).

Bitler, M.P., and J. Currie. 2005a. Does WIC work? The effects of WIC on pregnancy and birth outcomes. *Journal of Policy Analysis and Management* 24 (1): 73-91.

Bitler, M.P., and J. Currie. 2005b. The changing association between prenatal participation in WIC and birth outcomes in New York City: What does it mean? *Journal of Policy Analysis and Management* 24 (4): 687-90.

Bitler, M. P., J. Currie, and J. K. Scholz. 2003. WIC eligibility and participation. *Journal of Human Resources* 38 (4):1139-79.

Black, A.P., J. Brimblecombe, H. Eyles, P. Morris, H. Vally, and K. O'Dea. 2012. Food subsidy programs and the health and nutritional status of disadvantaged families in high income countries: A systematic review. *BMC Public Health* 12 (1): 1099.

Caliendo, M., and S. Kopeinig. 2008. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 22 (1): 31-72.

Currie, J., and I. Rajani. 2014. Within-mother estimates of the effects of WIC on birth outcomes in New York City. National Bureau of Economic Research. Working Paper 20400.

Currie, J., and M. Rossin-Slater. 2015. Early-life origins of life-cycle well-being: Research and policy implications. *Journal of Policy Analysis and Management* 34 (1): 208-42.

Devaney, B. 2010. WIC turns 35: Program effectiveness and future directions. In Reynolds, A., Rolnick, A., Englund, M., and Temple, J. (Eds.), *Childhood Programs and Practices in the First Decade of Life: A Human Capital Integration.* pp. 29-48. Cambridge: Cambridge University Press.

DiPrete, T.A., and M. Gangl. 2004. Assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological Methodology* 34 (1): 271-310.

Drichoutis, A.C., R. M. Nayga Jr, and P. Lazaridis. 2009. Can nutritional label use influence body weight outcomes? *Kyklos* 62 (4): 500-25.

Figlio, D., S. Hamersma, and J. Roth. 2009. Does prenatal WIC participation improve birth outcomes? New evidence from Florida. *Journal of Public Economics* 93 (1): 235-45.

Fox, M.K., W. Hamilton, and B. H. Lin. 2004. *Effects of Food Assistance and Nutrition Programs on Nutrition and Health: Volume 3, Literature Review*. USDA Economic Research Service, Food Assistance and Nutrition Research Report No. 19-3. Available https://www.ers.usda.gov/publications/pub-details/?pubid=46574 (accessed November 13, 2017).

Gai, Y., and L. Feng. 2012. Effects of federal nutrition program on birth outcomes. *Atlantic Economic Journal* 40 (1): 61-83.

Gordon, A., and L. Nelson 1995. *Characteristics and Outcomes of WIC Participants and Nonparticipants: Analysis of the 1988 National Maternal and Infant Health Survey.* Alexandria, VA: U.S. Dept. of Agriculture, Food and Nutrition Service, Office of Analysis and Evaluation.

Guenther, P. M., K. O. Casavale., J. Reedy, S. I. Kirkpatrick, H.A. Hiza, K. J. Kuczynski, L. L. Kahle, and S. M. Krebs-Smith. 2013. Update of the healthy eating index: HEI-2010. *Journal of the Academy of Nutrition and Dietetics* 113 (4): 569-80.

Hong, G., and B. Yu. 2008. Effects of kindergarten retention on children's social-emotional development: An application of propensity score method to multivariate, multilevel data. *Developmental Psychology* 44 (2): 407.

Hoynes, H., M. Page, and A.H. Stevens. 2011. Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program. *Journal of Public Economics* 95 (7-8): 813-27.

Joyce, T., D. Gibson, and S. Colman. 2005. The changing association between prenatal participation in WIC and birth outcomes in New York City. *Journal of Policy Analysis and Management* 24 (4): 661-85.

Keele, L. 2010. An overview of rbounds: An R package for Rosenbaum bounds sensitivity analysis with matched data. Available http://www.personal.psu.edu/ljk20/rbounds%20vignette.pdf (accessed November 9, 2017).

Kranz, S. and A.M. Siega-Riz. 2002. Sociodemographic determinants of added sugar intake in preschoolers 2 to 5 years old. *Journal of Pediatrics* 140(6):667-72

Morgan, R. 2015. *US Domestic Food Assistance Programs*. National Conference of State Legislatures. Published online at http://www.ncsl.org/research/human-services/u-s-domestic-food-assistance-programs.aspx (accessed November 2, 2017).

National Academies of Sciences, Engineering, and Medicine. 2017. *Review of WIC Food Packages: Improving Balance and Choice: Final Report.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/23655.

Oh, M., H.H. Jensen, and I. Rahkovsky. 2016. Did revisions to the WIC program affect household expenditures on whole grains? *Applied Economic Perspectives and Policy* 38 (4): 578-98.

Owen, A. L. and G.M. Owen. 1997. Twenty years of WIC: A review of some effects of the program. *Journal of the American Dietetic Association* 97(7):777-82.

Oliveira, V. and E. Frazão. 2015. *The WIC Program: Background, Trends, and Economic Issues, 2015 Edition*. Washington, DC: USDA Economic Research Service Economic Information Bulletin 134. Available https://www.ers.usda.gov/publications/pub-details/?pubid=43927 (accessed November 13, 2017).

Rosenbaum, P.R. 2002. *Observational Studies.* Springer.

Rosenbaum, P.R., and D.B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70 (1): 41-55.

Rosenbaum, P.R., and D.B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 39 (1): 33-8.

Rossin-Slater, M. 2013. WIC in your neighborhood: New evidence on the impacts of geographic access to clinics. *Journal of Public Economics* 102: 51-69.

Rubin, D.B. 1991. Practical implications of modes of statistical inference for causal effects and the critical role of the assignment mechanism. *Biometrics* 47 (4): 1213-34.

Sekhon, J.S. 2011. Multivariate and propensity score matching software with automated balance optimization: The Matching Package for R. *Journal of Statistical Software* 42(7):1-52. Available http://www.jstatsoft.org/v42/i07/. (accessed November 9, 2017).

Siega-Riz, A.M., S. Kranz, D. Blanchette, P.S. Haines, D.K. Guilkey, and B.M. Popkin. 2004. The effect of participation in the WIC program on preschoolers' diets. *The Journal of Pediatrics* 144 (2): 229-34.

Sonchak, L. 2016. The impact of WIC on birth outcomes: New evidence from South Carolina. *Maternal and Child Health Journal* 20 (7): 1518-1525.

Staff, J., M. E. Patrick, E. Loken, and J. L. Maggs. 2008. Teenage alcohol use and educational attainment. *Journal of Studies on Alcohol and Drugs* 69 (6): 848-58.

Swann, C. A. 2010. WIC eligibility and participation: The roles of changing policies, economic conditions, and demographics. *The BE Journal of Economic Analysis & Policy* 10 (1).

Thorn, B., C. Tadler, N. Huret, E. Ayo, C. Trippe, M. Mendelson, K.L. Patlan, G. Schwartz, and V Tran. 2015. *WIC Participant and Program Characteristics 2014.* Prepared by Insight Policy Research under Contract No. AG-3198-C-11-0010. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service. Available https://fns-prod.azureedge.net/sites/default/files/ops/WICPC2014.pdf (accessed October 29, 2017).

Tiehen, L. and E. Frazão. 2016. *Where Do WIC Participants Redeem Their Food Benefits? An Analysis of WIC Food Dollar Redemption Patterns by Store Type.* USDA Economic Research Service Economic Information Bulletin 152. Available https://www.ers.usda.gov/publications/pub-details/?pubid=44076 (accessed November 13, 2017)

Tiehen, L., and A. Jacknowitz. 2008. Why wait?: Examining delayed WIC participation among pregnant women. *Contemporary Economic Policy* 26 (4): 518-38.

Tiehen, L., and A. Jacknowitz. 2010. *WIC participation patterns: An Investigation of Delayed Entry & Early Exit.* USDA Economic Research Service Economic Research Report 109. Available https://www.ers.usda.gov/publications/pub-details/?pubid=44783 (accessed November 13, 2017).

United States Department of Agriculture Economic Research Service. 2017a. FoodAPS National Household Food Acquisition and Purchase Survey. Available https://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey.aspx (accessed November 13, 2017).

United States Department of Agriculture Economic Research Service. 2017b. Food Access Research Atlas: Documentation. Available https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation. (accessed November 5, 2017).

United States Department of Agriculture Economic Research Service. 2017c. Potatoes and Tomatoes are the Most Commonly Consumed Vegetables. Available: https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=58340 (accessed February 9, 2018)

United States Department of Agriculture Food and Nutrition Service. 2010. Dietary Guidelines for Americans 2010. Available https://www.cnpp.usda.gov/sites/default/files/dietary_guidelines_for_americans/PolicyDoc.pdf. (accessed November 9, 2017).

United States Department of Agriculture Food and Nutrition Service. 2016. WIC Fact Sheet. Available https://fns-prod.azureedge.net/sites/default/files/wic/WIC-Fact-Sheet.pdf. (accessed October 29, 2017).

Ver Ploeg, M., L. Mancino, J. E. Todd, D. M. Clay, and B. Scharadin. 2015. *Where do Americans Usually Shop for Food and How do they Travel to Get There?: Initial Findings from the National Household Food Acquisition and Purchase Survey*. USDA Economic Research Service Economic Information Bulletin 138. Available https://www.ers.usda.gov/publications/pub-details/?pubid=79791 (accessed November 13, 2017).

Wager, S. and S. Athey. 2017. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association* (forthcoming)

Wilde, P.E., P.E. McNamara, and C.K. Ranney. 1999. The effect of income and food programs on dietary quality: A seemingly unrelated regression analysis with error components. *American Journal of Agricultural Economics* 81 (4):959-71.

Wyse, A. E., V. Keesler, and B. Schneider. 2008. Assessing the effects of small school size on mathematics achievement: A propensity score-matching approach. *Teachers College Record* 110 (9): 1879-900.

Ye, Y., and L. A. Kaskutas. 2009. Using propensity scores to adjust for selection bias when assessing the effectiveness of alcoholics anonymous in observational studies. *Drug and Alcohol Dependence* 104 (1): 56-64.

Table 1. Descriptive Statistics for the Household (HH) Sample

Variable	Description	Eligible Non-Participants	WIC Participants		
			All	Redeemed WIC	Did Not Redeem WIC
		N = 505	N = 423	N=152	N=271
Healthy Purch. Index	Index (0 to 100)	50.388 (12.785)	50.259 (12.886)	55.958 (12.889)	47.062 (11.750)
Rural	Indicator (1 = rural tract)	0.236 (0.425)	0.234 (0.424)	0.257 (0.438)	0.221 (0.416)
Marital Status	Indicator (1 = married)	0.626 (0.484)	0.539 (0.499)	0.579 (0.495)	0.517 (0.501)
Hispanic	Proportion of HH members	0.247 (0.405)	0.392 (0.469)	0.351 (0.464)	0.415 (0.472)
African American	Proportion of HH members	0.156 (0.346)	0.172 (0.36)	0.112 (0.300)	0.205 (0.386)
Less High School	Indicator (1 = yes)	0.081 (0.273)	0.163 (0.37)	0.164 (0.372)	0.162 (0.369)
High School	Indicator (1 = yes)	0.196 (0.397)	0.324 (0.469)	0.309 (0.464)	0.332 (0.472)
Some College	Indicator (1 = yes)	0.384 (0.487)	0.369 (0.483)	0.355 (0.480)	0.376 (0.485)
College or Higher	Indicator (1 = yes)	0.339 (0.474)	0.144 (0.352)	0.171 (0.378)	0.129 (0.336)
Monthly Income	$1,000	5.144 (6.392)	2.834 (2.874)	3.071 (4.011)	2.702 (1.966)
WIC Eligible Children	Count	0.853 (0.569)	0.749 (0.718)	0.822 (0.790)	0.708 (0.672)
WIC Eligible Infants	Count	0.129 (0.335)	0.286 (0.483)	0.342 (0.529)	0.255 (0.453)
WIC Eligible Woman	Indicator (1 =yes)	0.158 (0.365)	0.296 (0.457)	0.283 (0.452)	0.303 (0.460)
WIC Clinic Access	Indicator (1 = access)	0.358 (0.480)	0.404 (0.491)	0.362 (0.482)	0.428 (0.496)
Supermarket Access	Indicator (1 = low access)	0.145 (0.352)	0.22 (0.415)	0.191 (0.394)	0.236 (0.426)
Self-reported Healthy Diet	Indicator (1 = healthy)	0.382 (0.486)	0.409 (0.492)	0.375 (0.486)	0.428 (0.496)

Educational attainment reflects the highest attainment of anyone in the household. Similarly, marital status = 1 if anyone in the household is married. Standard deviations are in parentheses.

Table 2. Logit model used to Match WIC Participants to Eligible Non-Participants.

	Estimate
Rural	0.063
	(0.044)
Marital Status	0.026
	(0.033)
Hispanic	0.161***
	(0.044)
African American	0.076
	(0.048)
Less High School	0.024
	(0.052)
Some College	-0.050
	(0.037)
College or Higher	-0.193***
	(0.046)
Monthly Income	-0.039***
	(0.007)
WIC Eligible Children	-0.022
	(0.024)
WIC Eligible Infants	0.157***
	(0.038)
WIC Eligible Woman	0.172***
	(0.037)
WIC Clinic Access	-0.025
	(0.035)
Number of Observations	928

Marginal effects (standard errors) are for the likelihood of WIC participation. Asterisks indicate significance: *, **, and *** at the 10, 5, and 1 percent levels, respectively. The model includes state fixed effects (not reported). See Table 1 for variable definitions.

Table 3. Effect of WIC Participation (Average Treatment Effect on the Treated (ATT)) on Healthy Purchasing Index Score from Matched WIC Subsamples.

	All Participants	Redeemed WIC	Did Not Redeem WIC
ATT estimate	2.742	9.443	-0.843
Standard error	1.351	1.573	1.421
p-value	0.042	<0.001	0.553
N (post-match)	534	178	337
Critical value	Sensitivity to Hidden Bias (Gamma)[a]		
0.05	1.13	2.82	-
0.10	1.17	3.05	-

[a] Magnitude of hidden bias (Gamma) from Rosenbaum sensitivity analysis required to change inference conclusions about the null hypotheses that ATT=0 from significant to insignificant at the given critical value.

Table 4. Average Treatment Effect on the Treated (ATT) Estimates from Matched Subsample of Households who Redeemed WIC by Component of the Healthy Purchasing Index (HPI).

				Hidden Bias (Gamma)[a]	
HPI Component	ATT Estimate	Standard Error	p-value	0.05	0.10
Adequacy components					
Total vegetables	-0.101	0.189	0.594	-	-
Greens and beans	0.237	0.254	0.350	-	-
Total fruit	0.798	0.202	<0.001	1.89	2.02
Whole fruit	0.534	0.247	0.031	1.30	1.39
Whole grains	1.829	0.338	<0.001	2.25	2.43
Total dairy	1.001	0.415	0.016	1.28	1.36
Total protein	0.076	0.179	0.673	-	-
Seafood and plant protein	0.901	0.261	<0.001	1.49	1.59
Fatty acids	0.398	0.473	0.400	-	-
Moderation components				-	-
Sodium	0.816	0.466	0.080	-	-
Refined grains	1.173	0.516	0.023	1.22	1.30
Empty calories	1.782	0.737	0.016	1.53	1.63

Higher values of each component indicate an improvement in nutritional quality of purchases. Reported p-values are for the null hypotheses that ATT = 0. Post-match N = 178.

[a] Magnitude of hidden bias (Gamma) from Rosenbaum sensitivity analysis required to change inference conclusions about the null hypotheses that ATT=0 from significant to insignificant at the given critical value.

Table 5. Effect of WIC Participation (Average Treatment Effect on the Treated (ATT)) on Healthy Purchasing Index Score from Matched WIC Subsamples: Excludes Households not Shopping at a WIC-approved Store.

	All Participants	Redeemed WIC	Did Not Redeem WIC
ATT estimate	2.671	9.385	-0.963
Standard error	1.353	1.579	1.425
p-value	0.048	<0.001	0.499
N (post-match)	527	178	330
Critical value	Sensitivity to Hidden Bias (Gamma)[a]		
0.05	1.13	2.80	-
0.10	1.17	3.02	-

[a] Magnitude of hidden bias (Gamma) from Rosenbaum sensitivity analysis required to change inference conclusions about the null hypotheses that ATT=0 from significant to insignificant at the given critical value.

Table 6. Effect of WIC Participation (Average Treatment Effect on the Treated (ATT)) on Healthy Purchasing Index Score from Matched WIC Subsamples: Healthy Purchasing Index Excludes Primary WIC Purchase Events.

	All Participants	Redeemed WIC	Did Not Redeem WIC
ATT estimate	-0.087	1.420	-0.843
Standard error	1.277	1.520	1.421
p-value	0.946	0.351	0.553
N (post-match)	534	178	337

Table 7 Effect of WIC Participation (ATT) on Healthy Purchasing Index Score from a Generalized Random Forest.

	ATT Estimate	Standard Error
All Participants	2.267***	0.820
Subsample that Redeemed WIC	7.720***	1.161
Subsample that Redeemed WIC (HPI exc. WIC events)	-0.109	1.105
Subsample that did not Redeem WIC	-0.903	0.873
HPI Component (Subsample that Redeemed WIC)[A]		
Adequacy components		
Total vegetables	-0.264*	0.135
Greens and beans	0.125	0.167
Total fruit	0.582***	0.143
Whole fruit	0.276	0.170
Whole grains	1.166***	0.258
Total dairy	0.978***	0.267
Total protein	-0.044	0.123
Seafood and plant protein	0.632***	0.182
Fatty acids	0.227	0.308
Moderation components		
Sodium	1.020***	0.332
Refined grains	0.496	0.337
Empty calories	2.445***	0.498

Asterisks *, **, and *** indicate statistical significance at the 10, 5, and 1 percent levels, respectively.

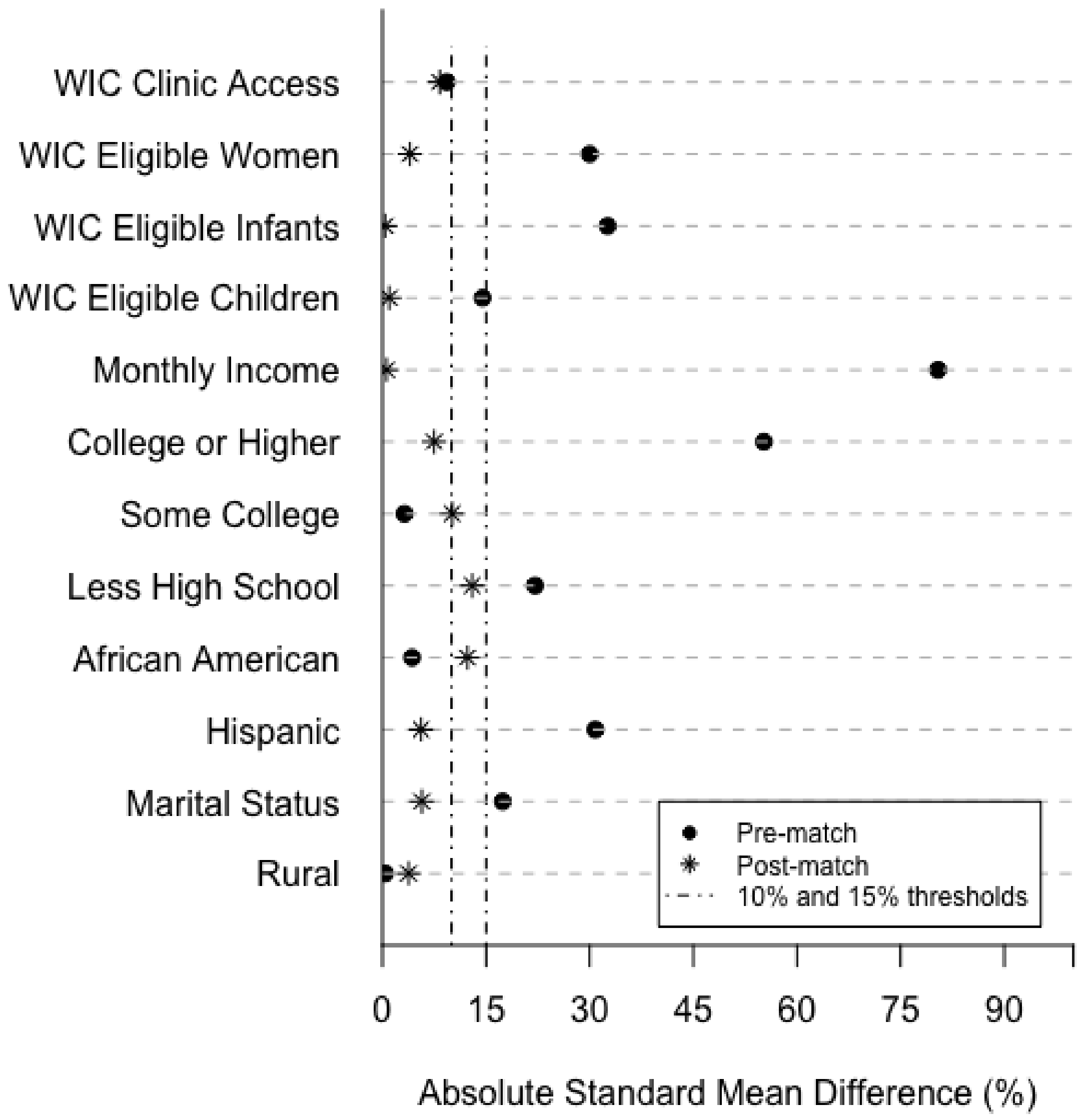

Figure 1. Pre and post-match balance comparisons by subsample.

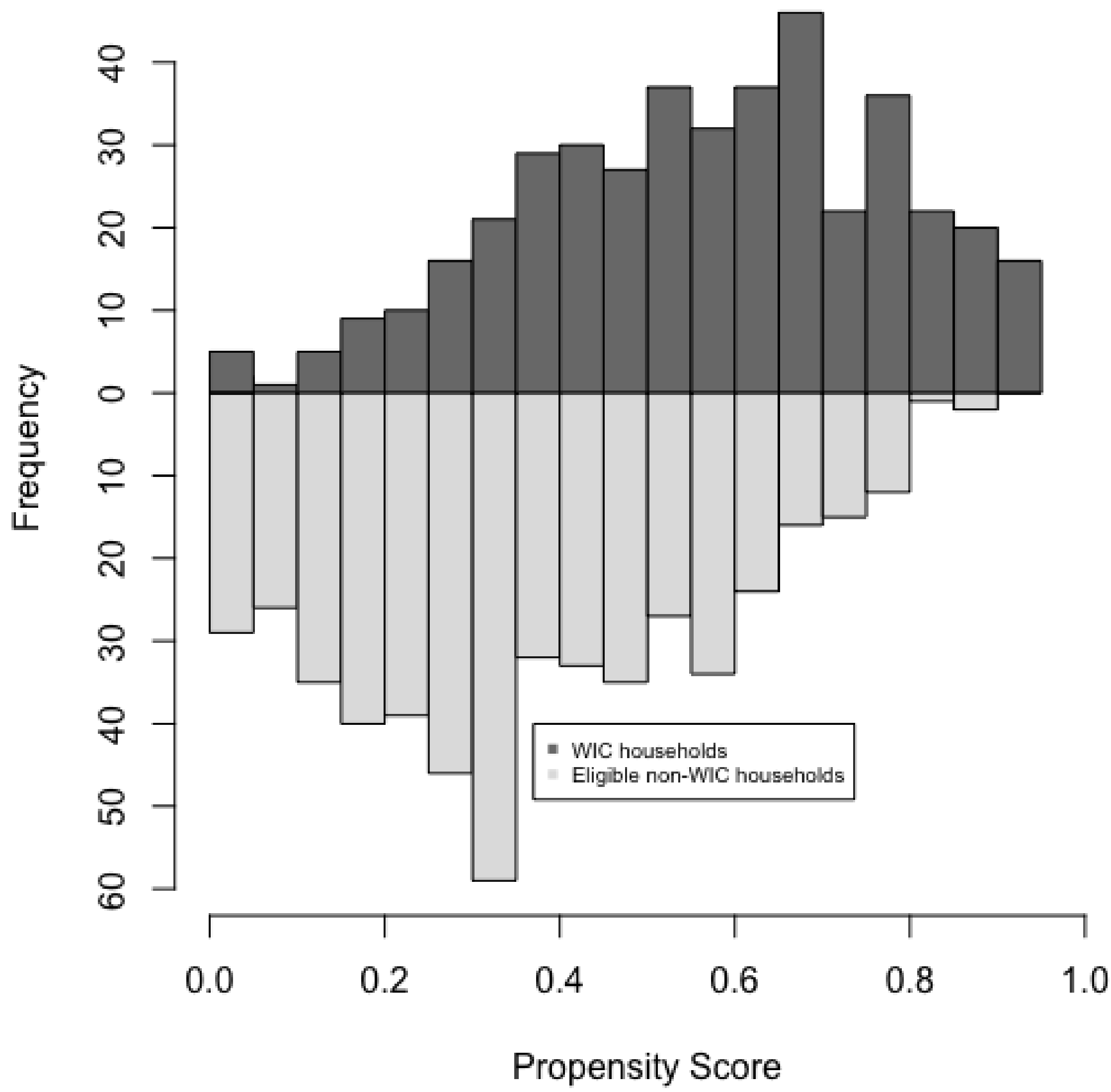

Figure 2. Distribution of propensity scores among WIC and eligible non-WIC households before the imposition of common support.

Appendix. Standard mean differences between WIC and eligible non-WIC households

Variable	Pre-match	Post-match
Rural	-0.378	3.802
Marital Status	-17.38***	5.732
Hispanic	30.783***	5.618
African American	4.325	-12.272*
Less High School	22.149***	-12.995*
Some College	-3.181	10.065
College or Higher	-55.273***	7.483
Monthly Income	-80.368***	0.642*
WIC Eligible Children	-14.496**	-1.043
WIC Eligible Infants	32.585***	0.392
WIC Eligible Woman	30.011***	4.037
WIC Clinic Access	9.33	8.292

Asterisks indicate significant differences based on a t-test of difference in means between the participants and eligible non-participants: *, **, and *** at the 10, 5, and 1 percent levels, respectively.