Chapter 4

# ANALYZING THE NUTRITIONAL EFFECTS OF THE EGYPTIAN FOOD SUBSIDY SYSTEM

The main hypothesis of this study is that Egypt's large food subsidy system has been ineffective in reducing undernutrition and may have contributed to the sustenance and even aggravation of both the double burden of malnutrition and the growth-nutrition disconnect. Accordingly, Egypt's exceptionalism in the global comparison with respect to these nutritional challenges may be partly driven by the country's food subsidy system, in combination with the three other drivers (the rapid nutrition transition, the succession of economic crises, and the persistent lack of nutrition-beneficial investments). Over time, the long-standing and expanding food subsidy system may have contributed to both nutritional challenges through two major effects. First, given its design until May 2014, the food subsidy system incentivized overconsumption of cheap, calorie-rich foods and unbalanced diets, which may have led to rapid increases in overweight and obesity and to slow or stagnant reduction of chronic child undernutrition. Second, given its heavy and growing burden to the public budget, the food subsidy system may have bound funds, which hence have been unavailable for more nutrition-beneficial investments, thereby maintaining and aggravating malnutrition indirectly. Thus, slow or stagnant reduction or even an increase of undernutrition may have contributed to a growing double burden of malnutrition if overnutrition rose faster than undernutrition declined, and to an aggravating growth-nutrition disconnect if it were accompanied by high economic growth.

In this chapter, we use impact evaluation methods and cross-sectional household survey data to explore whether the food subsidies affect child and maternal nutrition and the double burden of malnutrition in Egypt, as hypothesized. The results can provide evidence on the potential direct effect of the food subsidy system—the first of the aforementioned effects. The lack of suitable time-series data does not allow us to analyze the effects of the food subsidy system on people's nutritional status over time or to estimate the nutritional effects of past reforms. Instead, our analysis aims at identifying causal relationships between food subsidies and nutritional outcomes across

individuals and families. The existence of these relationships is fundamental for the hypothesized role of the food subsidy system as a driver of the double burden of malnutrition and the growth-nutrition disconnect.

We are particularly interested in the "dose-response relationship" between the acquired food subsidy levels and the nutritional status of the beneficiaries rather than only the nutritional effect of receiving any amount of food subsidies compared to receiving no food subsidies. We expect that the nutritional responses differ significantly between beneficiaries of high subsidy amounts and low subsidy amounts. Further, beneficiaries' nutritional status may become less or more responsive with increasing subsidy amounts, and an optimal subsidy level may exist; this suggests allowing the functional form of the dose-response relationship to be nonlinear and concave. Analyzing the hypothesized causal effects at different subsidy levels as well as relaxing the linearity assumption of the dose-response function may yield estimation results that have more accurate—and thus more useful—policy implications compared to an analysis that compares the causal effect only in a with-and-without-subsidies case or in a more-or-less-subsidies case.

## Methodology and Data

### Identification Strategy

Until May 2014, Egypt's food subsidies were issued through two separate programs with different eligibility criteria and different subsidy allotment criteria. The Baladi bread and flour subsidies were, in principle, accessible to every citizen in unrestricted amounts. In contrast, the subsidies under the ration card program were restricted to households with valid ration cards, and the quotas for subsidized rice, sugar, and cooking oil (and black tea) were allotted based on the number of household members registered on the ration card. Given these differences, we analyze the nutritional effects of the Baladi bread and flour program and the ration card program separately.

The nutritional effects of the subsidies under both programs cannot be simply assessed through an examination of the relationship between observed food subsidies and nutritional outcomes. Our analysis faces a common problem of observational studies for causal effects (Rosenbaum and Rubin 1983). When we estimate the effect of a treatment (such as the food subsidies) on an outcome (such as people's nutritional status) using observational data, a selection bias is likely, due to the nonrandom assignment of the treatment. A selection bias occurs when observed and unobserved characteristics of individuals

are associated with the probability of receiving the treatment—or the received amount of the treatment—and with the outcome. For example, an eligibility criterion for receiving a ration card in Egypt has been poverty, so low-income households are more likely to have ration cards than high-income households; low-income households also have less purchasing power for acquiring an expensive, well-balanced diet, which increases the risk of malnutrition among their household members. Similarly, Baladi bread outlets are geographically targeted to low-income neighborhoods, and Baladi flour was handed out only in some governorates in Upper Egypt at the time of the survey underlying our analysis.

The typical solution to the selection bias problem in the literature of experimental design is a randomized control trial (RCT), where random assignment to treatment balances observed and unobserved characteristics of individuals across treatment and comparison groups. Given that only one treatment level can be observed for each individual at a time, individuals of the comparison group who are similar to the treated individuals in everything but the treatment received are used as proxies for the counterfactual in RCTs. The randomization of treatment is generally impossible in observational studies, such as when evaluating social policies, because they are usually designed to be targeted at specific groups of individuals.

The second-best option is to mimic randomized assignment to treatment and comparison groups. As in many other observational studies, our choice of an impact evaluation method is constrained by data availability. Given that the 2010–2011 HIECS (CAPMAS and WFP 2011) is so far the only household survey that makes evaluating the nutritional effects of the Egyptian food subsidy system possible (Box 4.1), we are left with approaches that do not necessarily require a baseline or panel survey. Such quasi-experimental approaches include regression discontinuity design (RDD), instrumental variable (IV), and matching methods.

The RDD method can be used only for programs that have a continuous eligibility index with a clearly defined cutoff score to determine who is eligible and who is not (Gertler et al. 2011). Egypt's food subsidy programs do not meet this necessary condition, as our discussion of the food subsidy system in the previous chapter shows. Apart from that, the RDD method yields estimates of the program impact around the cutoff score—measuring local average treatment effects, which cannot necessarily be generalized to individuals with eligibility scores further away from the cutoff, such as high compliers. Considering the context of our study, we are less interested in such partial treatment effects.

An IV approach involves finding an instrument that is highly correlated with program participation but that is not correlated with the outcome—other than through program participation. We were unable to find such a credible instrument, mainly because (observed) household characteristics that determine participation in the food subsidy programs are naturally associated with nutrition, and thus the outcome variable, too. Even if there is an instrument that satisfies the exclusion restriction, it is likely not a good instrument, given the imperfect targeting of Egypt's food subsidies, especially of the ration card program. Because it has to account for unobserved heterogeneity, the instrument is very unlikely to be perfectly correlated with program participation, so only a subset of participants would be picked up by the instrument and resulting IV effect (Khandker, Koowal, and Samad 2010). Thus, at best we were able to estimate partial treatment effects, which, again, are of less interest in the context of this study.

Matching methods are suited for evaluating programs for which the rules of treatment assignment are less clear, but they require strong assumptions. We therefore decided to choose a matching approach and will demonstrate that the assumptions underlying our estimation model's specifications are reasonable. Matching methods use statistical techniques to construct a comparison group based on observed characteristics. Hence, they rely on the assumption that there is no unobserved, systematic difference between treatment and comparison groups that is associated with the outcome. We will detail and formalize this so-called "unconfoundedness assumption" in the next subsection. Because, by definition, it is impossible to statistically test for a selection bias stemming from unobservables, we note beforehand that the possibility of such a potential bias exists. Our strategy therefore is to minimize the probability of a potential selection bias as much as possible and to statistically assess how well the observed covariates explain program participation.

The propensity score matching (PSM) methodology offers a way to balance measured covariates across treatment and comparison groups, which helps to isolate the treatment effect. PSM also deals with the "curse of dimensionality" by compressing all relevant covariates into a single score—another attractive feature for our analysis. PSM was introduced by Rosenbaum and Rubin (1983) and has become a popular approach to estimate causal effects for cases of binary treatment, as in the case of program impact evaluations focusing on the effects of participation in the program compared to nonparticipation. However, many studies are rather concerned with the causal effects of a certain "dose" of a treatment—such as a received subsidy amount—relative to a lower or a higher "dose." Hirano and Imbens (2004) extended the

PSM methodology with binary treatment to cases where treatment is continuous. The generalization of the standard propensity score is referred to as the generalized propensity score (GPS), and the respective estimation function is known as the "dose-response function." In our analysis, we apply PSM methods with both binary and continuous treatments, which are explained in the next subsection.

Because receiving food subsidies under the ration card program is conditional on having a valid ration card, we first analyze whether participation in the ration card program has a causal effect on nutritional outcomes. For that, we apply the PSM method with binary treatment using samples of beneficiary and nonbeneficiary individuals/households. Then in order to go beyond the average treatment effect, we apply the PSM method with continuous treatment using samples restricted to individuals from households with ration cards, and adopt a quadratic specification of the dose-response function. In this second set of estimations, we hence analyze whether different amounts of received food subsidies cause different nutritional outcomes, while we allow the dose-response functions to take nonlinear and concave forms. However, for analyzing the causal effects of Baladi bread and flour subsidies on nutritional outcomes, we apply only the PSM method with continuous treatment, using samples of reported participating and nonparticipating individuals/households. The reason is that—unlike the ration card program—the access to subsidized Baladi bread and flour was de jure unrestricted at the time of the survey underlying our analysis. For both the ration card program and the Baladi bread and flour program, the amount of the allotted and utilized food subsidies depends on a set of household characteristics, which are also likely to determine the selection of households for the ration card program. The household characteristics included in the estimation models are described in the subsection after next.

We analyze the causal effects of the participation in the ration card program, the food subsidy amounts under the ration card program, and the food subsidy amounts under the Baladi bread and flour program on a large set of nutrition outcome indicators. They include indicators of chronic child undernutrition, child and maternal overnutrition, and the double burden of malnutrition at the individual and the family levels. For analyzing the effects on child undernutrition and child and maternal overnutrition, we use a continuous and a binary indicator each. Estimates for continuous indicators provide information about the effects across the nutrition spectrum, while estimates for binary indicators provide information about the probability of under- or overnutrition. Accordingly, child nutrition indicators are child HAZ and

child stunting (the respective binary indicator identifying chronic child undernutrition) as well as child BMIZ and child overweight (the respective binary indicator identifying child overnutrition). Similarly, maternal nutrition indicators are mother's BMI and maternal overweight—the respective binary indicator identifying maternal overnutrition. The binary indicator of the double burden of malnutrition at the individual level is child stunting and child overweight at the same time. The binary indicator of the double burden of malnutrition at the family level is child stunting in combination with maternal overweight. For complementarity, we also include a binary indicator of child overweight in combination with maternal overweight, as it may provide information about whether food subsidies contribute to overweight among children and their mothers from the same households at the same time.

Low child HAZ and child stunting is often caused by insufficient intake of absorbable micronutrients, especially zinc (Brown, Wuehler, and Peerson 2001; IZiNCG 2004). Dietary diversity is a strong predictor of micronutrient adequacy in children in developing countries, is associated with child HAZ, and is strongly associated with the micronutrient density of the diet (Ruel, Harris, and Cunningham 2013). Dietary diversity is typically measured by the number of different food groups consumed over a specified reference period (usually a maximum of seven days). The regular consumption of some food groups such as vegetables, legumes, meat and fish, and milk and dairy products is particularly important for nutrient adequacy. Several food items within these food groups are rich sources of absorbable zinc and other nutrients essential for children's physical growth. In order to explore whether unbalanced diets lacking diversity and frequency of nutritious foods qualify as a likely pathway through which food subsidies adversely affect child nutrition and hence drive the double burden of malnutrition—as hypothesized—we analyze the causal effects of the food subsidy programs on respective diet quality indicators. As dietary diversity indicator, we use a common food group count measurement—that is, the household dietary diversity score (HDDS). We complement it with measurements of the consumption frequency of four key food groups (vegetables, legumes, meat and fish, and milk and dairy products) over a one-week recall period. The indicators are described below.

We conduct our analysis of the causal effects of food subsidies on nutrition and diet quality outcomes for Egypt's urban and rural areas separately. Given that there are considerable urban-rural differences in the design and coverage of the food subsidy programs and in families' food consumption patterns (as discussed in the previous chapter), we expect that there are also structural

differences between urban and rural individuals in nutritional responses to received food subsidies. Moreover, the separate analysis may allow us to draw more specific conclusions from the estimation results.

## Estimation Framework

Every impact evaluation study has to overcome the basic problem of evaluation and thereby to deal with the possibility of selection bias. The basic problem is essentially a missing data problem. For assessing the effects of a policy or a program, we would like to know the difference in an outcome indicator with and without the policy or program—that is, the "treatment." However, we cannot observe both outcomes for the same individual at the same time. Comparing the mean outcomes of treated and untreated individuals is inappropriate, since the characteristics of treated and untreated individuals likely differ, even in the absence of treatment. This is because a policy or program is usually designed to target a specific group of individuals. Hence, in observational data, treatment assignment is not random. This can lead to a selection bias, where observed and unobserved characteristics of individuals are associated with the probability of receiving treatment and the outcome.

The matching approach is a possible solution to deal with the potential selection bias in observational data (Caliendo and Kopeinig 2008). The principal idea is to identify a group of untreated individuals who are similar to the treated individuals in all relevant pre-treatment characteristics so that they can serve as proxies for the counterfactual. Then, differences in outcomes between treated and untreated individuals can be attributed to the treatment. The underlying identifying assumption, as mentioned above, is known as "unconfoundedness" (Rosenbaum and Rubin 1983), "selection on observables" (Heckman and Robb 1985), or "conditional independence" (Lechner 2002). However, conditioning on all relevant covariates is limited in the case of a high-dimensional vector. As variables are added to the matching process, it becomes increasingly difficult to find exact matches of individuals with similar characteristics in both the treatment and comparison groups. Rosenbaum and Rubin (1983) suggest using propensity scores to deal with this "curse of dimensionality."

### PROPENSITY SCORE MATCHING WITH BINARY TREATMENT

Following Rosenbaum and Rubin (1983), we are interested in estimating the causal effect of a binary treatment on a binary or continuous outcome. Let the treatment, $T_i$, equal 1 if individual $i$ receives the treatment and 0 otherwise. The two potential outcomes are then $Y_i(1)$ and $Y_i(0)$ for each individual, $i$,

where $i = 1, \ldots, N$ and $N$ is the sample size. If $Y_i(1)$ and $Y_i(0)$ were observable, the treatment effect on individual $i$ would be directly observable as:

$$\tau_i = Y_i(1) - Y_i(0). \tag{1}$$

Since for each individual only one outcome is observable, estimating the individual treatment effect $\tau_i$ is impossible, and we have to concentrate on average treatment effects for groups of individuals. Two average treatment effects can be estimated. The first effect is the average treatment effect (ATE) on the total population:

$$\tau_{ATE} = E[Y(1) - Y(0)]. \tag{2}$$

The parameter gives the expected change in outcome if individuals in the population were randomly assigned to the treatment. The ATE is of little policy interest in general as well as of minor relevance in the context of our study, because it averages over individuals who might never be subject to treatment. The second, more prominent, average treatment effect is the average treatment effect on the treated (ATT):

$$\tau_{ATT} = E[Y(1) - Y(0) \mid T = 1] = E[Y(1) \mid T = 1] - E[Y(0) \mid T = 1]. \tag{3}$$

The parameter gives the expected change in outcome due to treatment for those individuals who were actually treated. Since the counterfactual mean for those being treated, $E[Y(0) \mid T = 1]$, cannot be observed, we need to identify an appropriate substitute for it. As noted above, using the mean outcome of untreated individuals, $E[Y(0) \mid T = 0]$, to estimate the ATT is usually no solution, because it is very likely that characteristics that determine the selection for receiving treatment also determine the outcome variable. Hence, even in the absence of treatment, the outcomes of individuals in the treatment and comparison groups would differ, leading to a selection bias. For the ATT, it can be noted as:

$$\begin{aligned} &E[Y(1) \mid T = 1] - E[Y(0) \mid T = 0] = \\ &\tau_{ATT} + E[Y(0) \mid T = 1] - E[Y(0) \mid T = 0], \end{aligned} \tag{4}$$

where the difference between $\tau_{ATT}$ and the term on the left-hand side of equation (4) is the selection bias. The parameter of the ATT is correctly estimated only if:

$$E[Y(0) \mid T = 1] - E[Y(0) \mid T = 0] = 0. \tag{5}$$

With observational data, some identifying assumptions are required to solve the selection problem, noted in equation (4). The unconfoundedness

assumption implies that systematic differences in outcomes between treated and untreated individuals with the same values for all relevant covariates can be attributed to the treatment. It ensures that the selection for receiving treatment satisfies some form of exogeneity. The unconfoundedness assumption can be written as:

$$Y(0), Y(1) \perp T \mid X, \tag{6a}$$

meaning that the potential outcomes $Y(0)$ and $Y(1)$ are independent of treatment assignment $T$, given a vector of covariates $X$ (that is, conditional independence). It implies that all variables that influence treatment assignment and potential outcomes simultaneously have to be observed, putting high requirements on data quality (Caliendo and Kopeinig 2008). Thus, unconfoundedness is technically a strong assumption.

Besides unconfoundedness, another required assumption is the "common support" or "overlap" condition. It can be written as:

$$0 < pr\,(T = 1 \mid X) < 1, \tag{7a}$$

meaning that individuals with the covariates $X$ have a positive probability of receiving treatment. The common support condition implies that there must be both treated and untreated individuals with each value of $X$, so that there is, for each treated individual, at least one untreated individual with the same characteristics. It rules out the phenomenon of perfect predictability of $T$ given $X$. Rosenbaum and Rubin (1983) refer to the unconfoundedness and common support conditions together as "strong ignorability." Under strong ignorability, the ATE (equation 2) and ATT (equation 3) can be determined for all values of $X$.

Since we are only interested in estimating the ATT, we can relax the assumptions as follows:

$$Y(0) \perp T \mid X \tag{6b}$$

and

$$pr\,(T = 1 \mid X) < 1. \tag{7b}$$

These assumptions of unconfoundedness for untreated individuals (equation 6b) and weak overlap (equation 7b) are sufficient for identification of equation (3), because the moments of the distribution of $Y(1)$ for the treated are directly estimable (Imbens 2004).

Conditioning on all relevant covariates is limited in the case of a high-dimensional vector $X$. For example, if $X$ has $m$ covariates and all are binary

variables, the total number of possible matches is $2^m$. To deal with this "curse of dimensionality," Rosenbaum and Rubin (1983) suggest using balancing scores. A balancing score, $b(X)$, is a function of the observed covariates, $X$, such that the conditional distribution of $X$ given $b(X)$ is the same for treated ($T = 1$) and untreated ($T = 0$) individuals. That is:

$$X \perp T \mid b(X), \tag{8}$$

Rosenbaum and Rubin (1983) show that, if potential outcomes ($Y(0)$, $Y(1)$) are independent of treatment assignment conditional on covariates, $X$ (equation 6a), they are also independent of treatment assignment conditional on a balancing score, $b(X)$. A possible balancing score is the propensity score. The propensity score is the probability for an individual to receive the treatment given a vector of observed covariates, $X$:

$$P(X) = pr(T = 1 \mid X). \tag{9}$$

Imbens (2004) shows that, if the unconfoundedness assumption holds, all biases due to observable covariates can be removed by conditioning solely on the propensity score. The condition of unconfoundedness given the propensity score can be noted as:

$$Y(0), Y(1) \perp T \mid P(X), \tag{10a}$$

and for estimating the ATT:

$$Y(0) \perp T \mid P(X). \tag{10b}$$

Given that the unconfoundedness assumption holds and that there is overlap between the treatment and comparison groups for all covariates, $X$, the PSM estimator for the ATT can be generally written as:

$$\tau_{ATT}^{PSM} = E_{(P(X) \mid T = 1)}\{E[Y(1) \mid T = 1, P(X)] - E[Y(0) \mid T = 0, P(X)]. \tag{11}$$

Thus, the PSM estimator is the mean difference in outcomes of treated and untreated individuals over the common support area, appropriately weighted by the propensity score distribution for the treated.

When implementing PSM, three choices have to be made. The first one concerns the model to be used for the estimation of the propensity score, the second one concerns the covariates to be included in this model, and the third one concerns the algorithm for matching treated and untreated individuals.

First, when estimating the probability of treatment versus nontreatment, logit and probit models usually produce similar results (Caliendo and Kopeinig 2008). We chose the logit distribution because it has more density

mass in the bounds, which tends to result in better fits of the data when sample sizes are large and extreme independent variable levels are present. Other than that, results from logit and probit models were found to be indistinguishable (Chambers and Cox 1967). It should also be noted that the role of the propensity score is only to reduce the dimensions of the conditioning covariates; as such, the propensity score has no behavioral assumption attached to it (Dehejia and Wahba 2002).

Second, our choice of the variables to be included in the propensity score model was guided by evidence from the food and health economics literature and our deep understanding of the Egyptian food subsidy system but limited by data availability and quality. In the next subsection, we discuss data issues and explain in detail the variables we included.

Third, we applied the *psmatch2.ado* routine developed by Leuven and Sianesi (2003) for application in Stata. We had to choose between several matching methods, including one-to-one (nearest neighbor or within caliper; with or without replacement), *k*-nearest neighbors, radius, kernel, local linear regression, spline-smoothing, and Mahalanobis matching. The choice of the matching algorithm involves trade-offs between bias and efficiency, especially in small samples (Caliendo and Kopeinig 2008). We chose the (Epanechnikov) kernel matching estimator, because—unlike nearest neighbor or radius matching estimators, for instance—it uses weighted averages of (nearly) all individuals in the comparison group for constructing the counterfactual outcome and therewith achieves low variance, because more information is used. Kernel matching can be seen as a weighted regression on an intercept with weight given by the kernel weights that vary with the point of evaluation (Smith and Todd 2005). The weights depend on the distance between each individual of the comparison group and the treated individual for which the counterfactual is being constructed. The average places higher weights on untreated individuals that are close to the treated individual in terms of the propensity score, and lower weights to more distant untreated individuals (if weights from a symmetric, non-negative, unimodal kernel such as the Epanechnikov kernel—the default kernel estimator in the *psmatch2.ado* routine—are used). The estimated intercept gives the estimate of the counterfactual mean. Applying kernel matching requires choosing the bandwidth parameter that involves finding a compromise between a small variance and an unbiased estimate of the true density function (Caliendo and Kopeinig 2008). Choosing a high bandwidth value yields a smoother density function, thus leading to a better fit and a decreasing variance between the estimated and true underlying density function. On the contrary, large bandwidth may

smooth away underlying features, leading to a biased estimate. We chose the default bandwidth of 0.06, which may optimize the trade-off between variance and bias (Heckman, Ichimura, and Todd 1997, 1998).

A drawback of nonparametric methods is that possibly individuals are used that are bad matches. Therefore, the proper imposition of the common support condition is critical. For estimating the ATT, it is sufficient that any combination of characteristics found in the treatment group may also be observed in the comparison group, whereas, for estimating the ATE, it is additionally required that any combination of characteristics seen in the comparison group may also be observed in the treatment group (Bryson, Dorsett, and Purdon 2002). We imposed the common support condition by enabling the respective option available in the *psmatch2.ado* routine. It drops treated individuals whose propensity score is above the maximum and below the minimum propensity score of untreated individuals. The number of observations that are "on-support" relative to those that are "off-support" indicates the area of common support.[1]

Finally, we test for the balancing property of the treatment and comparison groups after matching by calling the Stata routine *pstest.ado*—developed by Leuven and Sianesi (2012)—directly after *psmatch2.ado*. For each variable included in the model, the *pstest.ado* routine reports the means of the treatment and comparison groups and a standardized percentage bias, which is calculated as proposed by Rosenbaum and Rubin (1985), and performs t-tests for equality of the means.[2] It also provides overall measures of covariate balance between the treatment and comparison groups, including the pseudo-R-squared score (Sianesi 2004); the value of the likelihood-ratio test of the joint insignificance of all the regressors (LR chi-squared); the mean and median bias; and the Rubin's B and R scores (Rubin 2001). The pseudo-R-squared score indicates how well the regressors, $X$, explain the probability of selection into the treatment group (Sianesi 2004). There should be no systematic differences in the distribution of covariates between treatment and comparison groups after matching, and thus the pseudo-R-squared score should be fairly low (Caliendo and Kopeinig 2008). The likelihood-ratio test of the joint insignificance of all the regressors in the logistic regression model should not reject the null hypothesis of joint insignificance. Rubin (2001)

1 We report the numbers of "on-support" and "off-support" observations in the "Estimation Results" subsection.

2 Performing t-tests-based comparisons after PSM is controversial, because the t-test makes often untenable assumptions—including normal distribution of the covariates in treatment and comparison groups—and is sensitive to sample sizes (Austin 2009).

recommends that B should be less than 25 and R should be between 0.5 and 2 for the groups to be considered as sufficiently balanced.

#### PROPENSITY SCORE MATCHING WITH CONTINUOUS TREATMENT

Hirano and Imbens (2004) extend Rosenbaum and Rubin's (1983) PSM methodology for binary treatments to cases where treatment is continuous. The generalization of the propensity score for binary treatment is known as the GPS. Hirano and Imbens (2004) show that the GPS has a balancing property similar to the binary propensity score. Under certain conditions, the GPS removes all biases associated with differences in the covariates. The estimation function that indicates the relationship between a specific (continuous) treatment level—the "dose"—and its average outcome—the "response"—is referred to as the "dose-response function."

PSM with continuous treatment has been applied much less than PSM with binary treatment, given its fairly recent introduction. Applications of GPS matching include evaluations of the effects of maternal time on child development in the United States (Carneiro and Rodrigues 2009), breastfeeding duration on childhood obesity in the United States (Jiang and Foster 2013), South Africa's Child Support Grant on child nutrition (Aguëro, Carter, and Woolard 2006), Ethiopia's Productive Safety Net Program and related transfers on agricultural productivity (Berhane et al. 2014; Hoddinott et al. 2012), public subsidies on corporate research and development investments in China (Dai and Cheng 2015), and firms' export levels on their sales growth in Germany (Fryges 2009).

In line with the notation for deriving the binary PSM approach, let T be a continuous set of potential treatments defined over the interval $[t_0, t_1]$, and $Y(t)$ a set of potential outcomes for $t \in \mathrm{T}$. For each individual, $i$, where $i = 1, \ldots, N$ and $N$ is the sample size, we observe the actual treatment, $T_i$, the outcome corresponding to the treatment level received, $Y_i = Y_i(T_i)$, and a vector of treatment-unrelated covariates, $X_i$. We are interested in estimating the average dose-response function, $\mu(t) = E[Y_i(t)]$. For simplicity of notation, we omit the individual subscript $i$ in the sequel.

Following Hirano and Imbens (2004), we assume that $\{Y(t)\}_{t \in \mathrm{T}}$, $T$, and $X$ are defined on a common probability space; $T$ is continuously distributed with respect to the Lebesgue measure on T; and $Y = Y(T)$ is a well-defined random variable. The unconfoundedness assumption for the case of binary treatment (equation 6a) can be generalized to the case of continuous treatment as:

$$Y(t) \perp T \mid X \text{ for all } t \in \mathrm{T}. \tag{12}$$

This assumption is referred to as weak unconfoundedness, because it requires conditional independence only at each value of the observed treatment, $T$, instead of joint independence of all potential outcomes, $\{Y(t)\}_{t\in t_0,t_1}$.

Next, Hirano and Imbens (2004) define the propensity function as the conditional density of the treatment given the covariates, as

$$f_{T|X}(t|x) = r(t,x). \tag{13}$$

Then, the GPS is

$$R = r(T,X). \tag{14}$$

The GPS has a balancing property similar to that of the binary propensity score; that is, within Strata with the value of $r(t,x)$, the probability that $T = t$ does not depend on the value of $X$:

$$X \perp 1\{T = t\}|r(t,X). \tag{15}$$

Hirano and Imbens (2004) prove that, if assignment to treatment is weakly unconfounded given $X$, then it is weakly unconfounded given the GPS:

$$f_T\{t|r(t,X),Y(t)\} = f_T\{t|r(t,X)\} \text{ for every } t \in \mathrm{T}. \tag{16}$$

Based on that, they also prove that the GPS can be used to remove any bias associated with differences in the covariates. Finally, the dose-response function can be obtained in two steps:

$$\beta(t,r) = E[Y(t)|r(t,X) = r] = E[Y|T = t,R = r] \text{ and} \tag{17}$$

$$\mu(t) = E[\beta(t,r(t,X))]. \tag{18}$$

In the first step of the practical implementation, the conditional expectation of the outcome, $\beta(t,r) = E[Y|T = t,R = r]$, is estimated as a function of two scalar variables—the treatment level, $T$, and the GPS, $R$. In the second step, the dose-response function, $\mu(t) = E[\beta(t,r(t,X))]$, is estimated by averaging the estimated conditional expectation, $\hat{\beta}(t,r(t,X))$, over the GPS at each treatment level of interest.

In most previous economic applications of the GPS (e.g., Aguëro, Carter, and Woolard 2006; Berhane et al. 2014; Hoddinott et al. 2012), the treatment, $T$—or its transformation, $h(T)$—is assumed to be normally distributed, conditional on the covariates, $X$. However, this assumption may not always hold, as in our case. The Kolmogorov-Smirnov test for normality yields decisive evidence against the condition of normal distribution of the treatment

variable in all our estimation samples. Guardabascio and Ventura (2014) propose a flexible way to parametrically estimate the GPS when the treatment variable is not necessarily normally distributed. Hence, we applied the *glmdose .ado* routine developed by Guardabascio and Ventura (2014) for application in Stata.

Practically, the dose-response function is obtained in three steps (Guardabascio and Ventura 2014). First, the GPS is estimated for each individual given the received treatment level and the observed covariates (equation 14). Second, the conditional expectation of the individual outcome is estimated as a function of two scalar variables; these are the treatment level and the GPS (equation 17). Third, the dose-response function is estimated by averaging the estimated conditional expectation over the GPS at each treatment level of interest (equation 18). Guardabascio and Ventura's (2014) approach differs from previous approaches assuming normal distribution of the treatment variable (e.g., Bia and Mattei 2008) through the first step.

In detail, Guardabascio and Ventura (2014) replace the ordinary maximum likelihood estimator in the computation with the more flexible generalized linear model (GLM) estimator. By using the GLM estimator, the modeling differs from the ordinary regression by choosing the distribution of the treatment, $T$, from the exponential family—thus explicitly allowing non-normal distributions—and by applying a non-identity transformation of the mean of the treatment that is linearly related to the explanatory variables, $X$:

$$f(T) = c(T,\phi)exp\left\{\frac{T\theta - a(\theta)}{\phi}\right\} \text{ and} \tag{19}$$

$$g[E(T)] = \gamma' X. \tag{20}$$

The *glmdose.ado* routine requires specifying the distribution form of $a(\theta)$ and the functional form of $g(.)$. The choice of $a(\theta)$—referred to as the family—is guided by the nature of the treatment variable, because it determines the actual probability function. We chose the binomial distribution, which appears to be a good approximation of the treatment distribution in our estimation samples, as graph charts suggest. Irrespective of the distribution chosen, it holds for the first and the second moment that

$$E(T) = \dot{a}(\theta) \text{ and } Var(T) = \phi\ddot{a}(\theta), \tag{21}$$

where the dots symbolize the first and the second derivative with respect to $\theta$; and $\theta$ and $\phi$ are the canonical and the dispersion parameter, respectively. The choice of $g(.)$—referred to as the link function (which is monotonic and differentiable)—is suggested by the relationship between the treatment and the

explanatory variables, as it determines how the mean is related to the covariates $X$. We chose the logit link function, which is the default for the binomial family. The GPS estimator can be generally written as

$$\hat{R} = r(T,X) = c(T,\hat{\phi})exp\left\{\frac{T\hat{\theta} - a(\hat{\theta})}{\hat{\phi}}\right\}, \tag{22}$$

where $\hat{\theta}$ and $\hat{\phi}$ are the parameters of the chosen conditional distribution of the treatment that were estimated in a preceding step given the covariates.

For implementing the dose-response model, the *glmdose.ado* routine requires several data operations and model specifications, which we conducted consistently for all estimations. For the estimations related to the ration card program, we dropped households without ration cards from the sample, given the conditionality of having a ration card for receiving the subsidies.[3] In line with Hirano and Imbens (2004) and others (e.g., Aguëro, Carter, and Woolard 2006; Fryges 2009), we transformed the treatment variable into a fractional variable (after dropping outlier observations), such that the received subsidy level of a particular household is expressed as a fraction of the maximum subsidy level in the estimation sample. For each estimation sample, we divided the range of the received subsidies into three intervals, each of an equal number of observations (that is, tertiles), indicating low, moderate, and high subsidy levels. Consistent with previous studies (e.g., Berhane et al. 2014; Hirano and Imbens 2004; Hoddinott et al. 2012), we specified that the GPS is evaluated at the mean of the treatment variable within each treatment interval and that the GPS values are divided into quintiles for each treatment interval. Accordingly, for each of the covariates, the routine evaluates the balance by testing whether the mean in one of the three treatment groups is different from the mean in the other two treatment groups combined and adjusts the unbalanced covariates distribution for the GPS so that the balancing property is satisfied at a significance level lower than 1 percent. After the adjustment, the routine estimates the GPS using the GLM estimator, given the specified (binomial) distribution family and the specified (logit) link function for the treatment variable. Technically, we hence estimate a

3 Nonetheless, all samples include households with zero observations in the treatment variables. The treatment variables are described in the next subsection in detail. In the samples used for estimating the effects of the ration card program, there are a few households who hold ration cards but did not report consuming subsidized rice, sugar, or cooking oil during the 15-day food consumption recall period. As noted in the subsection "Characteristics of the Egyptian Food Subsidy System (under Study)" in the previous chapter, almost all households who have ration cards also use them. Less than 4 percent of the households in all urban samples and less than 2 percent in all rural samples underlying the dose-response model estimations for the ration card program report zero consumption for the subsidized foods.

fractional-logit (dose-response) model—a common case in empirical economics (Guardabascio and Ventura 2014).

We chose a quadratic estimation form for the dose-response function, because it allows for possible nonlinearity and concavity in the relationship between food subsidies and nutritional outcomes. We adopted a functional form that has both the GPS score and the fractional treatment variable in linear and quadratic terms as well as a term capturing interaction between the two variables. We determined the use of bootstrapping to derive standard errors and confidence intervals, performing 768 bootstrap replications. To determine a sufficiently large number of bootstrap replications, we used the *bssize.ado* routine developed by Poi (2004) for application in Stata. The routine implements Andrews and Burchinsky's (2000) method for choosing the number of bootstrap replications. The calculated number is the initial estimate of the number of bootstrap replications needed to obtain bootstrap standard errors that do not deviate by more than 5 percent from the ideal bootstrapped values with a probability of 95 percent. For defining the bandwidth, we chose the default option of the *glmdose.ado* routine. It applies an automatic procedure that estimates the unknown terms in the optimal global bandwidth as described in Fan and Gijbels (1996). Finally, we specified that the routine estimates the average potential outcome at 50 treatment levels that are equally spaced over the range of the observed treatment—that is, in steps of 2 percentage points in the fractional treatment level. The routine produces a graph of the dose-response function based on these estimates, which we present in the estimation results subsection.

### Survey Data and Estimation Variables

Our analysis of the causal effects of the Egyptian food subsidy system on nutritional outcomes uses data from the 2010–2011 round of the HIECS, conducted by Egypt's CAPMAS (CAPMAS and WFP 2011). This round provides us a unique opportunity to conduct such analysis, because it contained an additional module on household food security and the anthropometry of some household members. This module provides the data for the construction of all our nutrition and diet quality indicators. It was included upon request and with the support of WFP. No previous and—to date—no following round of the HIECS collected anthropometric measurements.

The basic HIECS is a large-scale, nationally representative household survey and includes a detailed household expenditure and consumption module that explicitly differentiates between subsidized food items and nonsubsidized food items. To the best of our knowledge, there is no other nationally

### BOX 4.1 Dilemma of the 2011 HIECS data: Household survey sampling

The HIECS is a household survey that collects detailed data on household incomes, expenditures, and consumption. It is a very large household survey, compared to comparable surveys in other countries. The 2010–2011 round collected data for 24,224 households (CAPMAS and WFP 2011). The HIECS is designed to be subnationally representative at the household level for all regions (Metropolitan areas, Upper Egypt, Lower Egypt, and Frontier Governorates), all governorates (there were 29 at the time of the 2010–2011 round), and urban and rural areas within the regions and governorates. The HIECS is carried out regularly—usually every other year—over a one-year period per round. The household sampling is designed to also yield representative estimates on a quarterly basis. Until 2010–2011, the HIECS was a purely cross-sectional survey (while part of the 2012–2013 round was designed as a panel of the 2010–2011 round), so changes in household variables cannot be tracked over time and variations between households need to be exploited for analysis.

CAPMAS has used its own local field survey teams for conducting paper and pencil interviewing (PAPI) for the HIECS, while another team at the headquarters has been responsible for digitalizing and cleaning the data, checking their consistency, requesting household revisits (if needed), and creating key variables. By default, households were revisited when absent during the first visit for interview. This system has been well established over decades. Nonetheless, it may be more prone to data collection and entry errors than a modern, well-established computer-assisted personal interview (CAPI)-based system. Moreover, traditional PAPI systems yield sufficiently large digital datasets for rigorous consistency checks usually only after several weeks or even months, which limits the possibility of timely reinterviewing of households for implausible records. Like in many other large household surveys, a possible consequence is that the HIECS data

representative survey in any other country with large food subsidies that provides both anthropometric measurements and detailed quantitative food subsidy information for the same households. However, the 2011 HIECS data also have some limitations, some of which present challenges to our analysis. We had to overcome data dilemmas related to the household survey sampling (Box 4.1), the construction of nutrition and diet quality indicators (Box 4.2), and the identification of food subsidy benefits (Box 4.3).

may contain a few flawed observations. For the econometric estimations, we therefore drop households/individuals that show outlier values in household income or anthropometric measurements.

For the first time and with financial and technical support from WFP, the basic HIECS in 2010–2011 was complemented with a module on household food security and anthropometry for children, adolescents, and women of reproductive age. The food security and nutrition module was carried out with a subsample of the full HIECS sample. Only those households that were visited in 2011 for the basic HIECS (during the third and fourth quarters of the one-year data collection period) were selected for the food security and nutrition survey; 11,802 households (48.7 percent of the full sample) completed both the basic HIECS and the complementary food security and nutrition survey. In 2011, the HIECS was carried out from January to June 2011 but suspended for most of February because of the political uprisings. CAPMAS and WFP deem the subsample to be representative at least nationwide and for Egypt's three main regions—the Metropolitan areas, Upper Egypt, and Lower Egypt—as well as for urban and rural areas nationally and regionally. CAPMAS granted us access to all expenditure and consumption data and data for several agreed-upon income variables from the basic HIECS for those households that were also interviewed for the food security and nutrition module. Hence, our analysis is based on this subsample. Our estimates from the basic HIECS data may therefore slightly deviate from estimates presented in official reports consulting the full sample. For the econometric estimations, we dropped households (and individuals) located in the sparsely populated desert Frontier Governorates (Matrouh, New Valley, Red Sea, North Sinai, and South Sinai), amounting to less than 1.5 percent of all valid observations in each estimation sample. We dropped these observations because of perfect predictability, given the set of covariates (which include households' residence by governorate and characteristics).

### ESTIMATION SAMPLES

All our estimations are based on three core datasets, which are (1) a child sample dataset, (2) a mother sample dataset, and (3) a household sample dataset. Because we ran the estimations for urban and rural areas separately, we split these core samples into urban and rural samples. All estimations of the effects on child nutrition employ the first datasets, all estimations of the effects on maternal nutrition employ the second datasets, and all estimations

### BOX 4.2 Dilemma of the 2011 HIECS data: Nutrition and dietary diversity indicators

The data collection for the complementary food security and nutrition survey was carried out by CAPMAS field survey teams (CAPMAS and WFP 2011). According to WFP and CAPMAS staff involved in the supervision of the data collection, enumerators were trained and received support in measuring child anthropometry from experienced health professionals. Nevertheless, anecdotal evidence suggests that taking adequate anthropometric measurements for infants and very young children appeared to be challenging in some cases. It was the first time in 2011 that the basic HIECS was complemented with an anthropometry module. Hence, the quality of the anthropometry data—especially for infants and very young children—may be somewhat lower than in specialized surveys, which prompted us to drop individuals with biologically implausible measurements from the samples.

Anthropometric measurements were taken for all children and adolescents between ages six months and 19 years and all women 20–49 years of age, who were present in the interviewed households when visited. Our analysis focuses on children ages 6–59 months and their nonpregnant mothers 20–49 years of age with biologically plausible anthropometric measurements. Identification of the children's mothers was not straightforward, because the household roster of the HIECS does not allow for explicitly recording the biological mother or caretaker. However, there is only one woman of reproductive age living in most households, whom we assume to be the biological mother or main caretaker. Only 9.8 percent of all households have more than one woman 20–49 years of age with plausible anthropometric measurements and also a child ages 6–59 months with plausible anthropometric measurements. In those cases, we employ a sequential procedure using combined information on a woman's age, marital status, and relationship to the household head to identify the likely mother (or female caretaker). We give the highest probability to married women of common

of the effects on household diet quality employ the third datasets. The estimations of the effects on the double burden of malnutrition at the family level use subsamples of the first and the second datasets—precisely with those child-mother pairs, where plausible anthropometric measurements are available for both individuals.

The child sample datasets include one child per household, who is 6–59 months old and has plausible HAZ and BMIZ values. For computing child HAZ and BMIZ from the height/length, weight, age, sex, and

childbearing age (that is, 20–28 years of age at the child's birth) who are the wives of the male household head and to women of common childbearing age who are the female household heads (who are often single mothers).

For deriving household food consumption and diet-related indicators, the 2011 HIECS data offer two principal sources. The first one is the food expenditure and consumption section of the basic HIECS. The second one is the food security section of the complementary food security and nutrition module. We use only data from the second source for computing our diet-related indicators for various reasons. The recall period of the food security section is 7 days, whereas it is 15 days for the food expenditure and consumption section. Seven days is a common—although already long—recall period for surveying dietary diversity. An even longer recall period is unsuitable for calculating established food-group-based dietary diversity indicators, because most households tend to consume foods from most considered food groups at least once over a long recall period. This is usually reflected in small cross-household variations of the indicators, which make them unsuited in (econometric) analysis. Choosing a dietary diversity indicator that counts the number of consumed food items instead of consumed food groups is not a real alternative because such an indicator does not provide information on whether households' diet is more diversified within one—perhaps relatively non-nutritious—food group or across food groups. Also, the food expenditure and consumption section does not allow us to calculate calorie and nutrient adequacy indicators because consumption quantities for processed foods—including bread, pasta, some dairy products, packed meat, and canned fish and vegetables—are not recorded. In contrast, the food security section provides data on the number of days a particular food group was consumed over the past 7 days. The listing of the food groups enables us to calculate the Household Dietary Diversity Score—a standard dietary diversity indicator. And the recorded number of days a particular food group is consumed allows us to explore the consumption frequency of key nutritious food groups in addition.

height measuring position records in the anthropometry section, we used the *zscore06.ado* routine, developed by Leroy (2011) for application in Stata. We dropped children with implausible anthropometric measurements from the samples, following standard WHO definitions of outlier measurements. Likely outliers are children who have a HAZ value below −6 or above +6 or a BMIZ value below −5 or above +5. If there were more than one child ages 6–59 months with plausible anthropometric measurements in the same

*Text continues on page 112*

### BOX 4.3 Dilemma of the 2011 HIECS data: Food subsidy benefits

The main purpose of the basic HIECS is to provide data for measuring poverty, inequality, and consumer price indexes. Hence, it is very detailed in household income and expenditure accounts, but it naturally falls short of what would be the ideal set of variables for our specific analysis (which can only be obtained from a specialized and probably very costly survey). We therefore need to compromise on optimal data adequacy in some aspects and rely more on econometric tools to demonstrate analytical robustness of our results.

The basic HIECS provides two alternatives to identify the subsidies that households receive under the ration card program. First, it gives the (self-) reported number of persons registered on the ration card, which can be used to calculate the allotted quotas of subsidized rice, sugar, and cooking oil (and tea). Second, the household food expenditure and consumption section reports the consumed quantities for the subsidized food items.

We decided against the first alternative, because there is a considerable possibility that a significant proportion of households systematically underreport the actual number of persons registered on the ration card. Households that have more persons registered on the ration card than are actually living in the household may underreport, because it earns them an illegally—or at least immorally—high benefit. Hence, they may be concerned that the provided information to the government's statistics organization may be used against them. Our HIECS data (CAPMAS and WFP 2011) suggest that systematic underreporting is indeed likely. According to the survey data, 72.4 percent of the population were beneficiaries of the ration card program in 2011. However, GASC—the ministerial department in charge of the food subsidy system—reported that 79.6 percent of the population were beneficiaries at the beginning of 2011, and 81.3 percent at the end of that year (CAPMAS 2014; MSIT 2014a, 2014b). The estimated number of household ration cards in the HIECS data is similar to (and even slightly higher than) the number from official sources.

Further, there is no accurate way to check if the allotted quotas derived from the reported number of registered persons are consistent with the consumed quantities of the subsidized foods, in such a way that the consumed quantities should not exceed the quotas. The comparison is inaccurate because the food consumption recall period is 15 days and the quotas are allotted on a monthly basis; the subsidized foods are nonperishable and therefore can be stored and then consumed in amounts above the monthly quotas; and the 15-day recall period can cut across two months, making purchases of twice the monthly quotas per recall period possible. Nevertheless, the correlation between the calorie amount of the monthly

quotas and the consumed calorie amount from the subsidized foods is high, with a coefficient of 0.518.

The variables to be included as covariates in the propensity score estimation models need to satisfy two principal conditions so that the models yield unbiased estimates. First, they should simultaneously influence the treatment variable and the outcome variable. Second, they should be unaffected by the treatment or the anticipation of it, which holds for variables that are fixed over time or measured before the treatment was received or its benefits are known. The limited availability of potentially deterministic variables in the HIECS data could interfere with the first condition, and the cross-sectional nature of the HIECS does not allow for verification of the second condition. Concerning the first condition, we would like to have had several specific variables that possibly better identify households' participation in the Baladi bread and flour program and the participation in and the received subsidy quotas under the ration card program.

For example, having georeferenced data on every household and on the Baladi bread outlet usually visited by the household, as well as the time spent queuing up for purchasing Baladi bread, would provide information on household transaction costs—a likely decisional factor of program participation and frequency of bread purchases. Related to the ration card program, we would like to have had copies of ration cards reporting the number of registered persons and quotas as well as information on when the household entered the program and on the eligibility criteria used to admit them to the program. Alternatively, more and more detailed information on certain household characteristics such as the year the household was established as well as (past) household affiliation with specific government and social institutions and organizations would have been desirable to check the characteristics' relevance for selection for the program. However, because such specific variables are unavailable from the HIECS data, we rely on proxy variables. To comply with the second condition as much as the data allow, we made several critical choices related to variable selection and specification. For example, for approximating household income levels, we used reported household income—instead of reported household expenditure (as used in the calculation of the official poverty estimates)—because reported household income is not influenced by the received food subsidies. Instead of using continuous variables for household income and household size, for instance, we used only categorical variables. Household income and household size are not fixed over a long period, ration cards seem not to have been revoked based on household incomes in the past, and household members have been continuously added to the cards of beneficiary households in past waves of program expansion. Household income quintiles and household size categories may be more time constant.

household, we randomly selected one child. We reduced the samples to one child per household to give equal weight to all sample households included in the estimations and hence to minimize a potential selection bias, considering that the HIECS is not designed to be representative for young children but households.

The mother sample datasets include the likely biological mothers or (female) caretakers of children ages 6–59 months who were selected for anthropometric measurement (including children with biologically implausible anthropometric measurements).[4] We dropped pregnant mothers because the BMIs yield incorrect indications of women's nutritional status during pregnancy. We identified mothers with biologically implausible BMIs using a common statistical definition of extreme outliers. We dropped any mother with a BMI that was above the third-quartile value in the core sample dataset (including mothers in both urban and rural areas) by a difference of three times the interquartile range (BMI = 44.79).[5] Like the child samples, the mother samples include only one individual per household.

The household sample datasets include all households with children 6–59 months old who were selected for anthropometric measurement (including children with biologically implausible anthropometric measurements).

Thus, all estimations are based on the same sets of individuals in the same households so that the estimates allow for drawing consistent conclusions across the different nutrition and diet quality indicators. The child datasets include 1,006 observations for the urban sample and 1,765 observations for the rural sample; the mother datasets include 1,140 observations for the urban sample and 1,823 observations for the rural sample; the child-mother pair datasets include 961 observations for the urban sample and 1,659 observations for the rural sample; and the household datasets include 1,130 observations for the urban sample and 1,911 observations for the rural sample.

### OUTCOME VARIABLES

We estimated the effects of the Egyptian food subsidy system on four child nutrition indicators, two maternal nutrition indicators, two child-mother nutrition indicators, one household diet diversity indicator, and four household food frequency indicators (Table 4.1). In detail, the child nutrition indicators are child HAZ and child stunting (the respective binary indicator

4 In the following, we omit "caretaker" for ease of readability.

5 The lower bound value, calculated as below the first-quartile value in the core sample dataset by a difference of three times the interquartile range, is below zero, so no observations are dropped due to that.

**TABLE 4.1** Overview of estimation model specifications with type of outcome and treatment variables

| | Ration card program | | Baladi bread & flour program |
|---|---|---|---|
| | PSM for program participation | Dose-response model: Subsidy level | Dose-response model: Subsidy level |
| **Chronic child undernutrition** | | | |
| Height-for-age z-score (HAZ) | continuous—binary | continuous—fractional | continuous—fractional |
| Stunting (HAZ < −2) | binary—binary | binary—fractional | binary—fractional |
| **Overnutrition of . . .** | | | |
| children | | | |
| Body-mass-index-for-age z-score (BMIZ) | continuous—binary | continuous—fractional | continuous—fractional |
| Overweight (BMIZ ≥ 2) | binary—binary | binary—fractional | binary—fractional |
| mothers | | | |
| Body mass index (BMI) | continuous—binary | continuous—fractional | continuous—fractional |
| Overweight (BMI ≥ 25) | binary—binary | binary—fractional | binary—fractional |
| **Double burden of malnutrition at . . .** | | | |
| individual level | | | |
| Child stunting & overweight (HAZ < −2 & BMIZ ≥ 2) | binary—binary | binary—fractional | binary—fractional |
| family level | | | |
| Child stunting (HAZ < −2) & maternal overweight (BMI ≥ 25) | binary—binary | binary—fractional | binary—fractional |
| **Overnutrition of children and their mothers** | | | |
| Child overweight (BMIZ ≥ 2) & maternal overweight (BMI ≥ 25) | binary—binary | binary—fractional | binary—fractional |
| **Household diet quality** | | | |
| Household dietary diversity score (HDDS) | categorical—binary | categorical—fractional | categorical—fractional |
| Consumption frequency of . . . | | | |
| vegetables | categorical—binary | categorical—fractional | categorical—fractional |
| legumes | categorical—binary | categorical—fractional | categorical—fractional |
| meat & fish | categorical—binary | categorical—fractional | categorical—fractional |
| milk & dairy products | categorical—binary | categorical—fractional | categorical—fractional |

**Source:** Authors' representation.

**Note:** PSM = propensity score matching.

identifying chronic undernutrition) as well as child BMIZ and child overweight (the respective binary indicator identifying child overnutrition). Maternal nutrition indicators are mother's BMI and maternal overweight (the respective binary indicator identifying maternal overnutrition). The binary indicator of the double burden of malnutrition at the individual level is child stunting and overweight in combination. The binary indicator of the double burden of malnutrition at the family level is child stunting in combination with maternal overweight. For complementarity, we also include a binary indicator of child overweight in combination with maternal overweight.

The chosen household dietary diversity indicator is the household dietary diversity score (HDDS), which was developed by the Food and Nutrition Technical Assistance project of the United States Agency for International Development (Swindale and Bilinsky 2006a, 2006b). The HDDS counts the number of food groups that the household consumed over the seven-day recall period of the food security section. It has a maximum score of 12 food groups (cereals, starchy roots, legumes, vegetables, fruits, meat and fish, eggs, milk and dairy products, sugars, edible fats and oils, condiments, and other meal additions). The household food frequency indicators count the number of days that the household consumed a particular food group over the seven-day recall period. We chose four food groups that are of particular importance for physical and mental child development and micronutrient adequacy in the Egyptian context, namely (1) vegetables, (2) legumes, (3) meat and fish (including beef, mutton, poultry, offal, fish, and seafood), and (4) milk and dairy products.

#### TREATMENT VARIABLES

We estimated the nutritional and dietary effects for the Baladi bread and flour program and the food ration card program separately but sought some sort of comparability of the estimation results for the two programs. To consistently sum up the acquired benefits from the different subsidized food items under each program, we needed to identify a common, nutrition-sensible denominator for constructing the treatment variables in the datasets for the dose-response model estimations. To meet this methodological need and to closely correspond to our main hypothesis, we expressed the treatment variables on a calorie consumption basis.

In the datasets for estimating the Baladi bread and flour program's effects, the treatment variable—before it is converted into a fractional variable—is the average calorie consumption amount per capita and day that the household obtained from the consumption of subsidized Baladi bread and Baladi bread

flour over the 15-day recall period of the food expenditure and consumption section in the basic HIECS (CAPMAS and WFP 2011). As for all processed foods, the HIECS data do not report the consumption quantity of Baladi bread but only the consumption expenditure, from which we derived the consumption quantity, using the fixed price of Baladi bread (EGP 0.05). The consumption quantities of Baladi bread flour and all foods subsidized under the ration card program are reported in the HIECS data. In the datasets for estimating the ration-card-program effects, the treatment variable—before converting it into a fractional variable—is the average calorie consumption amount per capita and day that the household obtained from the consumption of subsidized rice, sugar, and cooking oil over the 15-day recall period.[6]

In each estimation dataset, we first dropped households with outlier calorie consumption amounts that we defined as the observations above the 99-percentile value. Then, we converted the treatment variable into a fractional variable by dividing a household's calorie consumption amounts by the maximum calorie consumption amount in the dataset. Hence, the maximum value of the treatment variables is 1, and the minimum value is 0, indicating that—in the case of the ration card program—the particular household did not consume subsidized rice, sugar, or cooking oil, despite having a valid ration card. The binary treatment variable in the PSM estimations of the nutritional and dietary effects of participation in the ration card program takes values of 1 if the household has a valid ration card, and 0 otherwise. The household characteristics section of the HIECS consumption module questionnaire asked for this information explicitly.

#### INDEPENDENT VARIABLES

Our choice of the household characteristics variables to be included as covariates in the (binary and generalized) propensity score estimation models was guided by evidence from the food and health economics literature and our deep understanding of the Egyptian food subsidy system. We included only variables that simultaneously influence households' participation in the food subsidy programs or the acquired food subsidies and the nutrition outcome indicators. We also experimented extensively with alternative covariates, additional covariates, and the exact specification of the included covariates in preliminary work (not reported here). In all estimations, we used the same set

6 Black tea (without sugar) contains zero calories and is therefore excluded. Only 16.5 percent of all households with a valid ration card reported that they consumed subsidized black tea during the 15-day recall period (CAPMAS and WFP 2011).

of covariates—although the rationales for including some covariates differed across the estimations. This set includes categorical variables of household per capita income, household size, age of household head, and education level of household head as well as binary variables specifying whether the household is headed by a single woman; whether a household member receives a pension or social insurance or social assistance benefits; and whether a household member works in public administration, subsidy and social welfare services, education, health and civil service, or defense.[7] For executing the estimations, we transformed the categorical variables into binary variables, identifying income quintiles, five household size categories, and five age groups and four education levels for household heads.[8] The propensity score estimation equations also include binary variables identifying the location of the household by governorate (aggregate) and—for rural areas—a binary variable specifying whether the household farms agricultural land.[9]

As noted in the subsections describing the history of the Egyptian food subsidy system and its characteristics at the time of our survey data in the previous chapter, low household income status, single-female household head with dependents, pensioners, and beneficiaries of social insurance and social assistance programs were explicit criteria for household selection into the ration card program. Since the introduction of the ration card system, as well as throughout all past rounds of its expansion, households with low-income jobs and with economically and socially disadvantaged background were given highest priority. To the best of our knowledge, eligibility for the ration card program was usually defined through affiliation with a specific occupation or societal group. For example, households with members who were enrolled in public benefit programs or with the main income earner employed

7 Household income is the total net income of all household members as provided by CAPMAS. It includes wages and salaries, income from agricultural and nonagricultural activities—including value of own-consumption, income from financial and nonfinancial assets, value of cash and in-kind transfers, and imputed rent for owned housing.

8 Table A.11 in the Appendix presents the descriptive statistics of the estimation variables.

9 The urban sample includes 12 governorates or governorate aggregates. The rural sample includes 13 governorates or governorate aggregates. We aggregated some governorates because of insufficient observations by governorate. We aggregated governorates that were one governorate before 2010 (and meanwhile were partly consolidated again)—such as Giza and 6th of October Governorates, Luxor and Qena Governorates, and Cairo and (urban) Helwan Governorates—and that have similar demographic, ecological, and economic conditions, are located in the same region and bordering each other, and have similar food subsidy system coverage, such as the two Metropolitan governorates Port-Said and Suez; the urban areas of Kafr-Elsheikh and Gharbiya Governorates; the urban areas of Bni-Souef, Fayoum, and Menia Governorates; the rural areas of Domiyat and Dakahliya Governorates; the rural areas of Sharkiya and Ismailia Governorates; and the rural areas of Bni-Souef and Fayoum Governorates.

in the low-income public sector were among the primary beneficiaries. Moreover, public-sector employees in general are likely to be better connected than the rest of the population, which helped them to successfully apply for ration cards and get additional household members registered on the cards. In response to the food, fuel, and financial crises in 2007–2009, the ration card program was opened to all single female–headed households with dependents and other household groups with limited income-earning potential such as those who have no or low formal education. Therefore, we also account for household heads' attained education, differentiating between no formal education level and completed primary, secondary, and higher education.

Given the history and characteristics of the ration card program, we can expect that the longer a household has existed, the more likely it is that the household has managed to get into the ration card program; and the longer the household has been participating in the program, the more likely it is that the household has registered all (past) household members during one of the past expansion rounds. Because we have no data on the year when a household was established or the year when a beneficiary household entered the ration card program (Box 4.3), we use the age of the household head as a proxy variable, assuming that older household heads established their households earlier than younger ones.

We defined the five age categories in line with the duration between major modifications of the ration card program in the past. We grouped households whose head was 38 years or younger at the time of the HIECS in 2011, households whose head was older than 38 years and 48 years or younger, households whose head was older than 48 years and 58 years or younger, and households whose head was older than 58 years. A person 38 years of age in 2011 turned 18 years of age—the age of consent in Egypt—in 1991; that is, after automatic inclusion of newborns on the household ration cards was abolished. Still, this person was old enough to have founded a family and registered himself and his family during one of the rounds of program expansion, starting after 2005. A person 58 years of age in 2011 turned 18 years of age in 1971, that is, shortly after the beginning of Anwar Sadat's presidency and the rapid expansion of the ration card program in the 1970s. A household head 48 years of age in 2011 turned 18 years of age in 1981—the year of Sadat's assassination and the beginning of Hosni Mubarak's reign, so this man (or woman) was probably less likely to have been enrolled in the ration card program in his early adulthood as an independent household, given the tightening of the ration card program in the 1980s and 1990s.

Our rationale for the inclusion of household size as an independent variable and for the choice of household size categories follows a similar logic to that underlying the rationale for the household head age variable. Large households tend to incorporate more persons in different age groups than do small households. Containing a variety of age groups is likely to be associated with a higher probability of acquiring household ration cards and registering household members based on duration of household existence, whereas small households are more likely to have been recently established with a lower respective probability of acquiring cards and registering members. All else being equal, large households also have a higher probability of having at least one household member who fulfills the eligibility criteria for the ration card program. In combination with other variables included, the chosen household size categories allow for identification of certain household demographics that are likely to matter for selection into the ration card program and quota allotment. For example, a three-person household consisting of parents with one child is differentiated from a single-mother household with two children because of the variable identifying households headed by single women with dependents. Note that because of our estimation sample selection, all considered households have at least one young child (ages 6–59 months), so including a variable identifying two-person households leads to over-identification. As another example, a five-person household may consist of parents with their three children or grandparents with their daughter-in-law (or daughter) and her two children. Given patriarchal practices, the variables identifying the age category of household heads largely differentiate these two common cases.

Most Egyptian farm households are smallholder, semicommercialized farmers and thus may have been considered to be particularly vulnerable to food price and environmental shocks. These circumstances, in combination with the government's strong involvement in the agricultural sector in the past and the more recent (partial) agricultural liberalization, may have attracted particular attention to farmers as beneficiaries of public assistance (Ahmed et al. 2001). Until the mid-1980s, the production (and trade) of strategic food crops—mainly wheat, rice, and sugarcane, partly for supplying the food subsidy system—was centrally managed and fully under government control. Many Egyptian farmers are still organized in centrally established cooperatives, while membership in these cooperatives has possibly served as an eligibility criterion in rural areas.

All these variables are also likely to be associated with household participation in the Baladi bread and flour program. For example, poor households rely more on cheap Baladi bread than do rich households (Figure 3.4); single

mothers tend to have time constraints for getting to the bread outlets and queuing up because of alternative time needs for income earning and childcare, whereas pensioners and beneficiaries of social insurance and social assistance programs may have more spare time; and Baladi bread outlets may be located near public-sector institutions so that their employees have easy access to Baladi bread. Household size and age and education of the household head—who is the main decisionmaker in the household—are likely determinants of household resource allocation, including decisions on household food consumption, as numerous studies in the food economics literature have shown (e.g., Behrman and Deolalikar 1987; Cortez and Senauer 1996; Pitt and Rosenzweig 1985; Thomas 1990). In addition, long-established households—proxied by the households' age—tend to reside in old neighborhoods, which are more likely to have a high density of Baladi bread outlets than newly constructed neighborhoods are. And farm households may be located farther away from Baladi bread outlets (and Baladi flour warehouses), may be less mobile, and may face more time constraints for acquiring Baladi bread and flour than other occupational groups do.

The coverage of, and hence households' physical access to, the ration card program and the Baladi bread and flour program varies by governorate because of different demographics and infrastructure. It may also vary because of differently influential representation in policy and administration, among other reasons. And Baladi bread flour was handed out in only some governorates at the time of the 2010–2011 HIECS.

All independent variables included in the propensity score estimation equations are also likely to affect household diet quality and individual nutritional outcomes, as evidence from the food and health economics literature suggests. Rich households can afford pricier, more nutritious food and better healthcare than poor households. More nutritious food and better healthcare are associated with better diet quality and lower risks of micronutrient deficiencies and chronic undernutrition (e.g., Brooks-Gunn and Duncan 1997; Ecker and Qaim 2011; Pinstrup-Andersen and Caicedo 1978; Ravallion 1990; Wolfe and Behrman 1983). Yet as discussed before in great detail, higher income does not necessarily lead to better nutrition.

The sex of the household head—as the main decisionmaker in the household—matters for the nutritional status of household members and for young children in particular, because it accounts for common gender differences in nutrition-relevant decisionmaking (e.g., Handa 1994; Kennedy and Peters 1992; Onyango, Tucker, and T. Eisemon 1994; Rashid, Smith, and Rahman 2011; Thiele, Mensink, and Beitz 2004). However, including just the sex of

the household head as a covariate yields misleading results given the Egyptian context. According to our HIECS data (CAPMAS and WFP 2011), most female-headed households in Egypt (79.2 percent) are households where no adult male is present, so decision-making power is with her by necessity, whereas most households with adult males are headed by men (95.8 percent). Therefore, we included instead a variable that identifies households headed by single mothers because single mothers usually face major time constraints due to their often dual responsibility as main income earners and main child caretakers (and often even main meal preparers), which may adversely affect child nutrition.

Evidence from the health literature suggests that formal education is a strong predictor of nutrition—especially parental education for the nutritional status of young children (e.g., Semba et al. 2008; Strauss 1990; Thomas 1990). Given the Egyptian context, we included the education level of the—usually male—household head.[10] Similarly and as mentioned above, the age of the household head may influence decisionmaking on the allocation of household resources, and the household size matters for nutritional effects associated with (reverse) household economies of scale (Deaton and Paxson 1998; Lanjouw and Ravallion 1995; Nelson 1988) and sharing of childcare responsibilities (Doan and Bisharat 1990; Popkin 1980).

Households who receive social benefits or engage in public-sector activities may also have better access to nutrition-relevant public services, including primary healthcare. Farm households often have access to nutritious own-produce, such as vegetables, legumes, and animal-source foods but are also frequently exposed to parasites in (polluted) irrigation water, giving rise to secondary malnutrition (e.g., Berti, Krasevec, and FitzGerald 2004; Carletto et al. 2015; Sibhatu, Krishna, and Qaim 2015; Stephenson, Latham, and Ottesen 2000). Moreover, farming is a physical labor–intensive activity, so farmers have a lower risk of overweight due to high calorie expenditure. A household's location is likely to affect food consumption and individual nutrition due to the local infrastructure endowment and quality, such as consumer marketplaces, water and sanitation systems, and healthcare facilities, among others.

10 Alternative estimation specifications having the attained education level of the mother yield similar—though less strong—results (not reported here).

### POTENTIAL ESTIMATION BIASES

As mentioned above, our identification strategy rests on the assumption that there are no other covariates than the ones considered that simultaneously influence treatment and outcome variables. This is a strong assumption for any empirical analysis of our kind, given survey data limitations. To make this assumption most plausible, we chose the empirical specifications to be consistent with established evidence from the food and health economics literature and our deep understanding of the Egyptian food subsidy system.

While we are confident that the chosen specifications of the estimation models exploited the full data potential, it is still conceivable that there are unobserved covariates that influence the treatment variables and the outcome variables at the same time. This could lead to biased estimates because the propensity score cannot adjust for the unobserved (systematic) differences between the treatment and comparison groups. Such a potential bias due to unobservables would arise if, for example, malnourished households manipulate their ration-card-program eligibility regarding the number of registered persons, whereas well-nourished households abstain from the possibility of this fraud. We doubt that this case is a likely source of a potential bias, however, as it assumes that only/predominantly malnourished households have some sort of (unobserved) means to successfully manipulate the selection process.

It should be clear that a potential bias due to unobservables arises only if the unobserved covariates influence both the treatment variable and the outcome variable. We acknowledge that there are likely several unobserved covariates that explain the selection into the ration card program and the acquired subsidy level under the ration card program and the Baladi bread and flour program, but we also believe that the probability that these covariates systematically influence the nutritional outcomes is fairly small. There are also variables that determine nutritional outcomes but are unlikely to influence the treatment variable. An example is children's age. Child stunting mainly evolves over the first two to three life years. However, neither the age of the children nor their number per household should influence households' participation in the ration card program or the received subsidy quotas because all children in our analysis are younger than five years of age, and the last expansion of the ration card program was before all of their births. Thus, we believe that the possibility of a bias due to unobservables is fairly small, and even if there is such an unobserved covariate, we believe that the potential bias does not compromise our main findings.

**TABLE 4.2** Overview of PSM estimation results: Direction of causal effects

| | Ration card program | | | | Baladi bread & flour program | |
|---|---|---|---|---|---|---|
| | PSM for program participation[a] | | Dose-response model: Subsidy level[b] | | Dose-response model: Subsidy level[b] | |
| | Urban | Rural | Urban | Rural | Urban | Rural |
| **Chronic child undernutrition** | | | | | | |
| Height-for-age z-score (HAZ) | **+** | + | – | **0** | 0 | – |
| Stunting (HAZ < –2) | – | + | + | **+/–** | 0 | **0** |
| **Overnutrition of . . .** | | | | | | |
| children | | | | | | |
| Body-mass-index-for-age z-score (BMIZ) | + | + | **+** | 0 | –/+ | + |
| Overweight (BMIZ ≥ 2) | 0 | + | **+** | **+/–** | –/+ | + |
| mothers | | | | | | |
| Body mass index (BMI) | + | 0 | **–/+** | 0 | + | 0 |
| Overweight (BMI ≥ 25) | **+** | + | **+** | 0 | + | 0 |
| **Double burden of malnutrition at . . .** | | | | | | |
| individual level | | | | | | |
| Child stunting & overweight (HAZ < –2 & BMIZ ≥ 2) | – | + | **+** | **+/–** | –/+ | **0** |
| family level | | | | | | |
| Child stunting (HAZ < –2) & maternal overweight (BMI ≥ 25) | – | + | +/– | **–/+** | +/– | **0** |
| **Overnutrition of children and their mothers** | | | | | | |
| Child overweight (BMIZ ≥ 2) & maternal overweight (BMI ≥ 25) | + | + | + | +/– | **–/+** | 0 |

## Estimation Results

This section presents our estimation results of the likely effects of the ration card program and the Baladi bread and flour program on individual nutrition and household diet quality among households with children ages 6–59 months—referred to as "families." Table 4.2 provides an overview of all results, indicating the direction of the estimated causal effects. First, we discuss the estimation results of the PSM with binary treatment that we used to assess the nutritional and dietary effects of participation in the ration card program compared to nonparticipation. Then, we discuss the estimation results of the PSM with continuous treatment for the ration card program. The results provide evidence on the hypothesized dose-response relationships

| | Ration card program | | | | Baladi bread & flour program | |
|---|---|---|---|---|---|---|
| | PSM for program participation[a] | | Dose-response model: Subsidy level[b] | | Dose-response model: Subsidy level[b] | |
| | Urban | Rural | Urban | Rural | Urban | Rural |
| **Household diet quality** | | | | | | |
| Household dietary diversity score (HDDS) | 0 | – | **–/+** | –/+ | + | **+/–** |
| Consumption frequency of . . . | | | | | | |
| vegetables | – | + | **+** | **–/+** | + | 0 |
| legumes | + | – | **–** | 0 | **+/–** | 0 |
| meat & fish | – | – | – | **–/+** | + | **–** |
| milk & dairy products | – | **–** | **–** | **–/+** | **–/+** | **–/+** |

**Source:** Authors' representation.

**Note:** PSM = propensity score matching.

[a] The signs indicate that the nutrition or dietary outcome indicator is higher (+) or lower (–), at the means, in the treatment group than in the comparison group. The 0 indicates that the mean difference is less than 1 percentage point. Bold signs indicate that goodness of fit is sufficient and an estimated coefficient of at least one subsidy-level variable is statistically significant at 10%. Normal, non-bold signs indicate that goodness of fit is sufficient but no coefficient of the subsidy-level variables is statistically significant at 10%. Bold gray signs indicate goodness of fit is insufficient and an estimated coefficient of at least one subsidy-level variable is statistically significant at 10%. Non-bold gray signs indicate goodness of fit is insufficient.

[b] The signs indicate that the nutritional or diet quality indicator continuously increases (+) or decreases (–) with higher subsidy levels, as suggested by the slope of the estimated dose-response function curve. The +/– and –/+ indicate that the estimated dose-response function follows an inverted or upright bell-shaped curve. The 0 indicates that the estimated dose-response function has a flat curve with no apparent slope. For defining the shape of the dose-response function curve, its confidence intervals are taken into consideration. Bold signs indicate that goodness of fit is sufficient and an estimated coefficient of at least one subsidy-level variable is statistically significant at 10%. Normal, non-bold signs indicate that goodness of fit is sufficient but no coefficient of the subsidy-level variables is statistically significant at 10%. Gray signs indicate goodness of fit is insufficient.

between the acquired food subsidy level (expressed on a calorie consumption basis and in relative terms) and nutrition and diet quality indicators among beneficiaries of the ration card program (identified by the possession of ration cards). Finally, we discuss the estimation results of the PSM with continuous treatment for the Baladi bread and flour program. The results provide evidence on the hypothesized dose-response relationships between the Baladi bread and flour subsidies and nutrition and diet quality outcomes among all families (considering that participation in the Baladi bread and flour program is unrestricted). Since the coefficients of the estimated dose-response functions have no direct interpretation, we present the estimated dose-response functions in graphical form.

### Effects of Ration-Card-Program Participation

The goodness of fit of all the logistic regression models for the estimation of the binary propensity scores is decent (with pseudo R-squared scores between 0.22 and 0.25), and it is slightly higher for the rural samples than for the urban samples.[11] For all PSM estimations, the area of common support between the treatment group (that is, ration cardholders) and the comparison group (that is, non–ration-card holders) is fairly large. Imposing the common support condition drops 11–14 percent of the total number of observations in the urban samples and 9–11 percent of the total number of observations in the rural samples (Tables 4.3–4.6). The test statistics for the balancing property of the treatment and comparison groups after matching suggest that there is no substantial overall covariate imbalance in all logistic regression models. The overall biases tend to be slightly higher for the rural samples than for the urban samples.[12] As required, the pseudo R-squared scores are low, and the likelihood-ratio test statistics do not reject the null hypothesis of joint insignificance of the regressors. The mean biases for each variable included in the models are usually quite small and never exceed 10.5 percentage points. The highest mean biases occur due to matching on household size in the urban samples and the household head's education level in the rural samples. The estimated overall mean biases and overall median biases never exceed 4 percentage points and 3 percentage points, respectively. Rubin's B-value never exceeds 25, and Rubin's R-value is always between 0.5 and 2. Accordingly, the groups of the treated and untreated individuals/families can be considered as sufficiently balanced in all estimation models.

Overall, our PSM estimations with binary treatment provide ambiguous evidence of the causal effects of ration-card-program participation on individual nutrition and household diet quality. This finding is in line with our expectations, considering that there are large differences in household subsidy levels (and low subsidy levels are unlikely to create a significant nutritional response), so averaging over all families with ration cards might not be meaningful. Nevertheless, the ATT estimates indicate a few significant effects and some notable tendencies.

In urban areas, children from families with ration cards have on average significantly higher HAZs than their peers from families without ration cards, and child stunting tends to be less prevalent among ration cardholders,

11 Tables A.12–A.15 in the Appendix present the logistic regression results of the binary propensity score estimations.

12 Tables A.16–A.19 in the Appendix present the balancing property test statistics.

**TABLE 4.3** Estimated ATT of ration-card-program participation on child nutrition

| | Urban | | | | | Rural | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treated | Untreated | Diff. | Std. err. | t-stat | Treated | Untreated | Diff. | Std. err. | t-stat |
| HAZ | –0.861 | –1.186 | 0.325 | 0.164 | 1.98 | –0.930 | –0.957 | 0.027 | 0.147 | 0.18 |
| Stunting | 0.271 | 0.320 | –0.049 | 0.038 | –1.30 | 0.262 | 0.251 | 0.011 | 0.034 | 0.34 |
| BMIZ | 1.055 | 0.877 | 0.177 | 0.147 | 1.21 | 1.030 | 0.922 | 0.108 | 0.132 | 0.82 |
| Overweight | 0.268 | 0.270 | –0.002 | 0.038 | –0.05 | 0.285 | 0.243 | 0.042 | 0.035 | 1.20 |
| Stunting & overweight | 0.139 | 0.159 | –0.021 | 0.030 | –0.70 | 0.121 | 0.119 | 0.002 | 0.026 | 0.09 |
| Common support (observations) | | | | | | | | | | |
| Total | 390 | 616 | | | | 1,082 | 683 | | | |
| On-support | 339 | 616 | | | | 977 | 683 | | | |

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** ATT = average treatment effect on the treated; BMIZ = body-mass-index-for-age z-score; HAZ = height-for-age z-score.

**TABLE 4.4** Estimated ATT of ration-card-program participation on maternal nutrition

| | Urban | | | | | Rural | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treated | Untreated | Diff. | Std. err. | t-stat | Treated | Untreated | Diff. | Std. err. | t-stat |
| BMI | 29.68 | 29.02 | 0.66 | 0.42 | 1.59 | 28.45 | 28.23 | 0.22 | 0.37 | 0.60 |
| Overweight | 0.832 | 0.758 | 0.074 | 0.033 | 2.28 | 0.713 | 0.736 | –0.024 | 0.034 | –0.70 |
| Common support (observations) | | | | | | | | | | |
| Total | 445 | 695 | | | | 1,127 | 696 | | | |
| On-support | 394 | 695 | | | | 999 | 696 | | | |

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** ATT = average treatment effect on the treated; BMI = body mass index.

while the mean difference is statistically insignificant (at common levels) (Table 4.3). These children tend to also have higher average BMIZs, but there is no indication for an increased probability of overweight. In rural areas, the average BMIZ and the probability of overweight tend to be higher among children from beneficiary families than among those from nonbeneficiary families. The average BMI of mothers of young children tends to be higher among urban beneficiaries than urban nonbeneficiaries, and, consistent with this pattern, overweight among these mothers is significantly more likely (Table 4.4). There is no clear evidence for average effects of ration-card-program participation on maternal overnutrition in rural areas. Thus, on average, household participation in the ration card program seems to contribute to better nutrition among young children but also leads to overnutrition among

**TABLE 4.5** Estimated ATT of ration-card-program participation on the double burden of malnutrition and the coexistence of child and maternal overnutrition at the family level

| | Urban | | | | | Rural | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treated | Untreated | Diff. | Std. err. | t-stat | Treated | Untreated | Diff. | Std. err. | t-stat |
| Child stunting & maternal overweight | 0.237 | 0.278 | –0.041 | 0.036 | –1.14 | 0.176 | 0.174 | 0.002 | 0.031 | 0.07 |
| Child overweight & maternal overweight | 0.246 | 0.217 | 0.029 | 0.037 | 0.78 | 0.206 | 0.169 | 0.037 | 0.033 | 1.13 |
| Common support (observations) | | | | | | | | | | |
| Total | 367 | 594 | | | | 1,028 | 631 | | | |
| On-support | 317 | 594 | | | | 914 | 631 | | | |

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** ATT = average treatment effect on the treated.

**TABLE 4.6** Estimated ATT of ration-card-program participation on household diet quality

| | Urban | | | | | Rural | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Treated | Untreated | Diff. | Std. err. | t-stat | Treated | Untreated | Diff. | Std. err. | t-stat |
| HDDS | 10.58 | 10.68 | –0.09 | 0.09 | –1.05 | 10.35 | 10.45 | –0.11 | 0.08 | –1.34 |
| Consumption frequency of | | | | | | | | | | |
| vegetables | 3.148 | 3.337 | –0.188 | 0.163 | –1.15 | 3.577 | 3.504 | 0.072 | 0.150 | 0.48 |
| legumes | 4.460 | 4.401 | 0.059 | 0.192 | 0.31 | 4.222 | 4.437 | –0.215 | 0.166 | –1.30 |
| meat & fish | 2.761 | 2.971 | –0.210 | 0.138 | –1.52 | 2.548 | 2.582 | –0.035 | 0.110 | –0.32 |
| milk & dairy products | 4.701 | 4.771 | –0.070 | 0.179 | –0.39 | 4.311 | 4.723 | –0.411 | 0.166 | –2.48 |
| Common support (observations) | | | | | | | | | | |
| Total | 453 | 677 | | | | 1,171 | 740 | | | |
| On-support | 398 | 677 | | | | 1,063 | 740 | | | |

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** ATT = average treatment effect on the treated; HDDS = household dietary diversity score.

mothers of young children in urban areas, while it seems to contribute to child overnutrition without increasing maternal overnutrition notably in rural areas. Nevertheless, finding a family where the child and her mother are overweight is more likely among beneficiaries than nonbeneficiaries in both urban and rural areas (Table 4.5). However, the mean differences are statistically insignificant. The double burden of malnutrition at the family level tends to be less common among beneficiaries than among nonbeneficiaries in urban

areas, while there is only a tiny (insignificant) difference in rural areas. The probabilities of the double burden of malnutrition and the coexistence of child and maternal overnutrition at the family level are considerably higher among both beneficiary and nonbeneficiary families in urban areas than families in rural areas.

Household diet quality tends to be somewhat lower overall among families with ration cards than families without ration cards in urban and rural areas. In particular, nutritious food groups tend to be consumed less frequently among beneficiary families; this is especially true of vegetables and meat and fish in urban areas, and of legumes and milk and dairy products in rural areas (Table 4.6). Yet the mean difference between beneficiaries and nonbeneficiaries is statistically significant only for the consumption frequency of milk and dairy products in rural areas. There is virtually no difference in household dietary diversity between beneficiaries and nonbeneficiaries in urban and rural areas. The mean differences vary around just 1 percent.

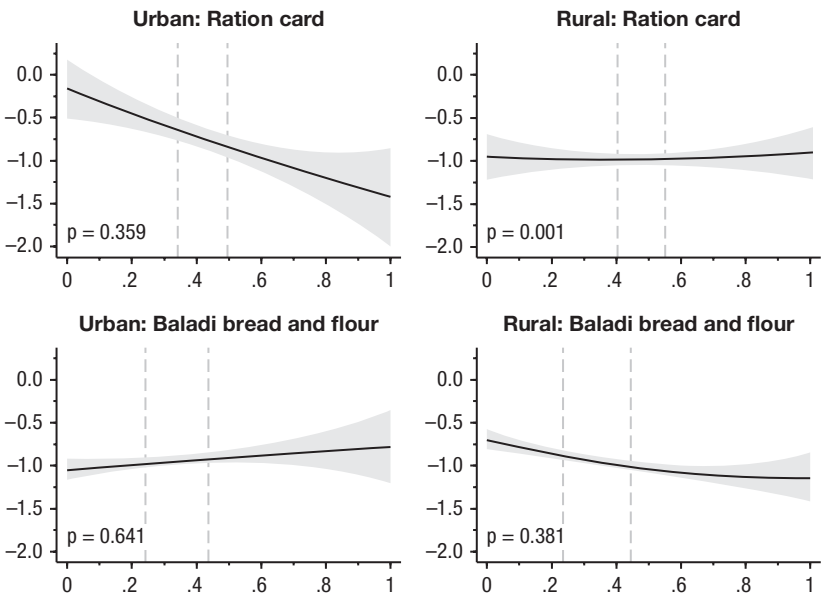
### Effects of Ration-Card-Program Subsidies

The goodness of fit of all GLM estimates of the GPSs—in the first step of the dose-response models for assessing the effects of the subsidy levels of the ration card program among beneficiary families—is satisfactory (with Pearson scores between 0.11 and 0.15).[13] In the second step, the goodness of fit of the estimated dose-response functions is generally low (with R-squared and pseudo R-squared scores of less than 0.09).[14] However, low goodness-of-fit measures are not uncommon in cross-sectional regression models that explore variations in anthropometric measurements of individuals—especially for height-based indicators of child nutrition (e.g., Breisinger and Ecker 2014; Christiaensen and Alderman 2004), mainly due to large genetic variations and common measurement problems. Large heterogeneity in observed variables in combination with relatively small sample sizes possibly also contributes to low goodness-of-fit measures and statistically insignificant coefficient estimates in the second step of our dose-response models, especially in the case of the estimations of subsidy-level effects of the ration card program in urban areas. Nevertheless, our estimation results indicate consistent tendencies for several nutrition and diet quality indicators.

13 Tables A.20–A.23 in the Appendix show the GLM results of the GPS estimations for the ration-card-program subsidies.

14 Tables A.24–A.37 in the Appendix show the dose-response function estimates for the effects of ration-card-program subsidies.

**FIGURE 4.1** Dose-response functions for child HAZ



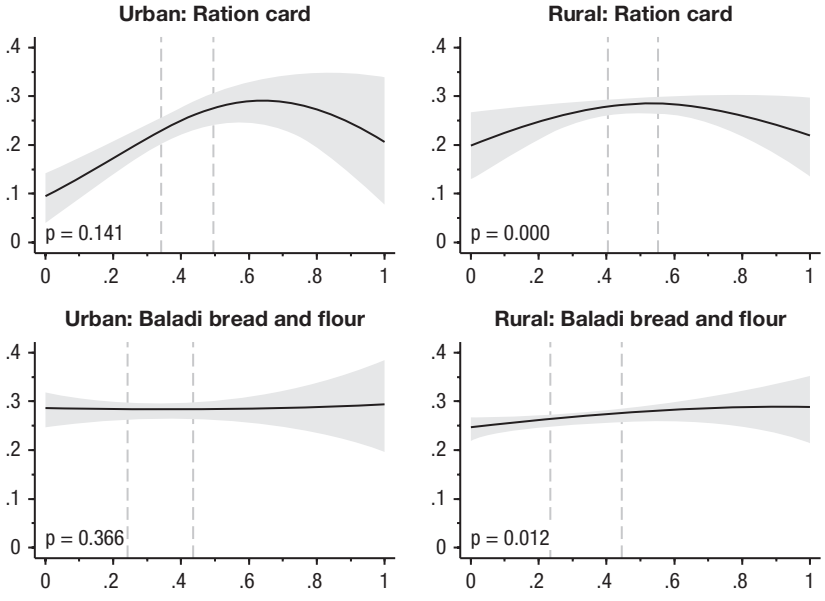

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate child height-for-age z-scores (HAZs), and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

In urban areas, the HAZs of children from families with ration cards tend to sharply decrease with increasing ration card subsidy levels in a linear manner, and the probability of child stunting tends to continuously increase across children from families with subsidy levels—other than very high subsidy levels (Figure 4.1 and Figure 4.2). Hence, the risk of chronic child undernutrition tends to increase with the consumed calories from subsidized rice, sugar, and cooking oil among urban beneficiary families. However, the average risk of chronic child undernutrition tends to be lower among urban beneficiaries compared to urban nonbeneficiaries, as the ATT estimates above suggest. Thus, it is the received subsidy amounts (above average) that tend to negatively affect children's nutritional status—but not household participation in the ration card program per se. Although the continuous and binary indicators of chronic undernutrition suggest a consistent tendency among ration cardholders in urban areas, it should be noted that the coefficient estimates in both

**FIGURE 4.2** Dose-response functions for child stunting




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

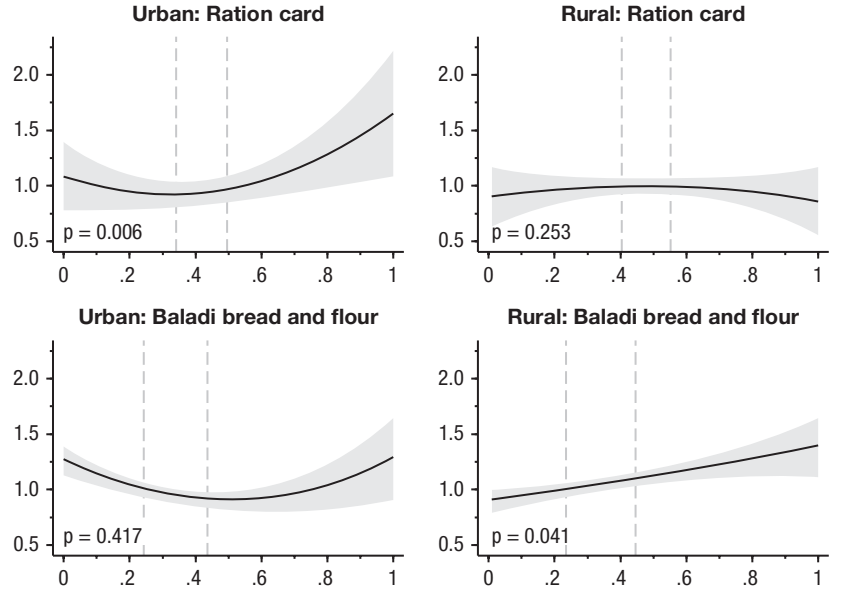
**Note:** The y-axes indicate the probabilities of children being stunted, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > chi-sq.). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

dose-response functions are statistically insignificant at common levels, and their goodness-of-fit measures are insufficient.

For rural areas, the dose-response functions indicate that the HAZs of children from beneficiary families and their probability of being stunted are quite unresponsive to different ration-card-program subsidy levels, with a slight—though statistically significant—tendency toward a higher child stunting probability at medium subsidy levels (Figure 4.1 and Figure 4.2). Differences in the responsiveness of nutrition indicators between urban and rural beneficiaries may be explained by various factors. A possible explanation for the differential effects on child HAZ and stunting is that children's physical growth in rural areas seems to be predominantly influenced by factors other than the acquired ration card subsidies and particularly by those factors that are accounted for by the GPS, such as infrastructural conditions at household locations. The coefficient estimates of the GPS (linear and quadratic)

**FIGURE 4.3** Dose-response functions for child BMIZ




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** BMIZ = body-mass-index-for-age z-score.

The y-axes indicate child BMIZs, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

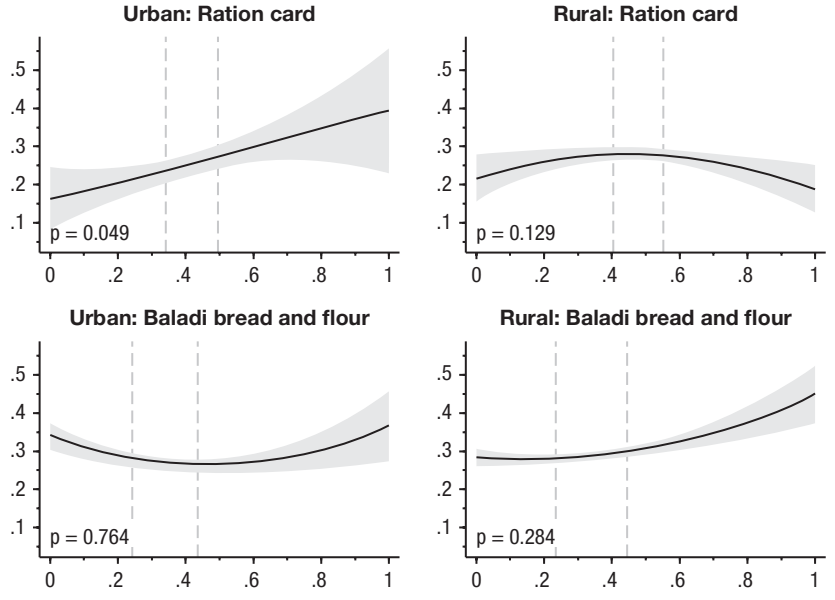
*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

variables in the dose-response functions for child HAZ and child stunting in rural areas are large and highly statistically significant (at the 1 percent level)—unlike those in the dose-response functions for child HAZ and child stunting in urban areas. And the GPS estimations for child nutrition in urban and rural areas show the largest differences in the coefficient estimates for several binary variables identifying the governorate where the beneficiary families reside.[15]

The BMIZs of children from urban beneficiary families increases with increasing ration-card-program subsidy levels, at least for children from families with medium and high subsidy levels (Figure 4.3). This estimation result is consistent with the ATT estimate above, according to which children from urban beneficiary families tend to have higher average BMIZ than children

15 See Tables A.20, A.24, and A.25 in the Appendix.

**FIGURE 4.4** Dose-response functions for child overweight




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate the probabilities of children being overweight, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > chi-sq.). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

from urban nonbeneficiary families. Among urban beneficiaries, the probability of child overweight increases with increasing subsidy levels in a linear manner and at high margins (Figure 4.4). The goodness of fit of the dose-response functions for child BMIZ and child overweight is sufficient, and the coefficient estimates of the subsidy-level variables—which define the shape of the curves—are statistically significant, providing strong evidence for the effects of ration-card-program subsidies on child overnutrition in urban areas.[16]

This is different for the dose-response functions of the subsidy-level effects on child BMIZ and child overweight in rural areas. They have insufficient goodness-of-fit measures and—in the case of the estimation for the effect on child BMIZ—statistically insignificant coefficient estimates.[17] The

16 See Tables A.26 and A.27 in the Appendix.

17 See Tables A.26 and A.27 in the Appendix.

dose-response model estimation for the effect on child overweight indicates a slightly higher probability of child overweight at medium subsidy levels. Like chronic child undernutrition, child overnutrition seems to be more responsive to ration-card-program subsidies in urban areas than in rural areas. Possible explanations here are related to differences in household availability of subsidized calories and individual calorie expenditures. Rural families have, on average, lower calorie consumption from subsidized rice, sugar, and cooking oil than urban households do, and there are fewer families with very high subsidy levels in rural areas than in urban areas, given that the per capita quotas of rural beneficiary households exceed the number of actual household members less often than among urban households (Table 3.4).[18] In addition, children in rural areas may be more physically active than their peers in urban areas and burn off relatively more surplus calories, including those ingested with the subsidized foods.

The dose-response function of the subsidy-level effect of the ration card program on the BMIs of mothers from urban beneficiary families follows an inverted, flattened bell-shaped curve, implying that mothers' BMIs tend to be lowest at medium subsidy levels (Figure 4.5). However, it should be noted that the confidence interval expands considerably toward both ends of the curve and that the shape of the curve is essentially determined by the interaction between the subsidy level and the GPS variable, indicating that there is large heterogeneity at low and high subsidy levels and that the effect at a particular subsidy level distinctly varies subject to the covariates.[19] Moreover, the average BMI at the lowest point of the curve equals 29.8, while a BMI of 30 or more identifies obesity—the extreme form of overweight. Hence, a possible interpretation of this result is that, above a certain level of overnutrition, additional cheap calories from subsidized foods may not lead to markedly more weight gains. Investigating the effect on nutritional status relative to a cutoff level may be more revealing in such a case.

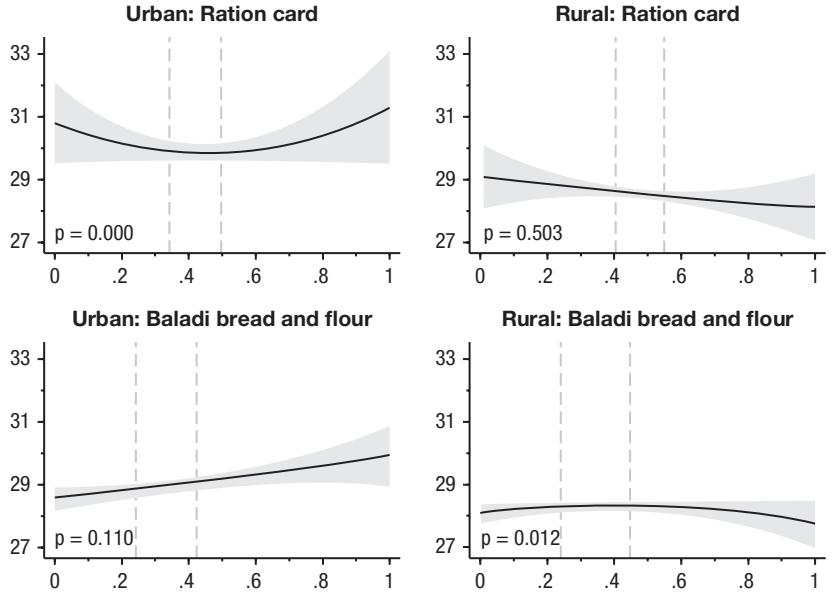
Accordingly, the dose-response model estimation for maternal overweight among urban families shows that the probability of maternal overweight in urban areas increases with higher subsidy levels at large—but declining—margins (Figure 4.6). The dose-response function has sufficient goodness of fit, and the respective coefficient estimates of the subsidy-level variables are highly statistically significant.[20] Mothers from urban families with medium

18 See Table A.11 in the Appendix.

19 See Table A.28 in the Appendix.

20 See Table A.29 in the Appendix.

**FIGURE 4.5** Dose-response functions for mother's BMI




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

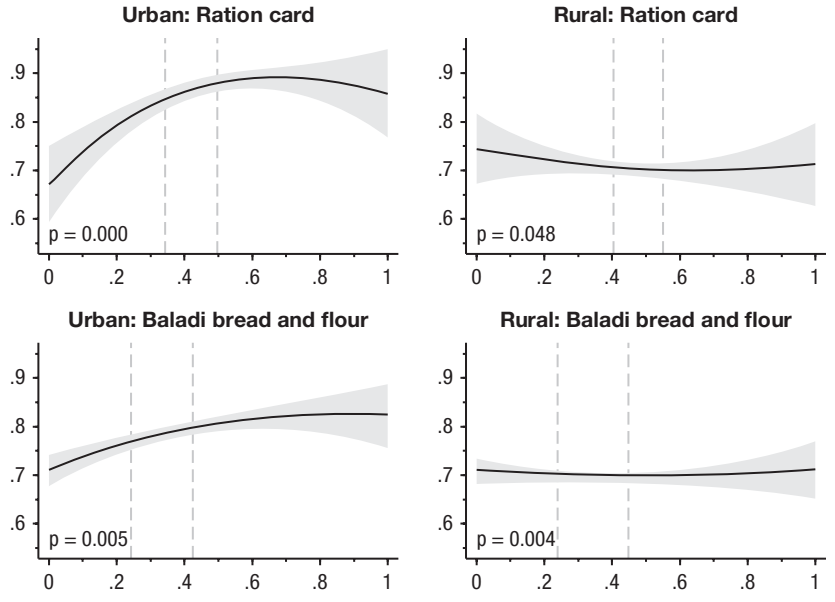
**Note:** The y-axes indicate mothers' body mass indexes (BMIs), and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

or high subsidy levels have a risk of overweight that is on average more than 10 percentage points higher than the risk of mothers from urban families with low subsidy levels. Consistent with that, the respective ATT estimates above indicate that the average risk of maternal overweight is also significantly higher among urban beneficiaries than among urban nonbeneficiaries. Thus, we found strong evidence that the ration card program contributes to maternal overweight, in addition to child overweight, in urban areas.

The dose-response model estimations of the subsidy-level effects on mother's BMI and maternal overweight among rural beneficiary families yield coefficient estimates of the subsidy-level variables that are highly statistically insignificant, and the goodness of fit of the dose-response function for mother's BMI is far from being sufficient (Figure 4.5 and Figure 4.6). The seeming unresponsiveness of the maternal overnutrition indicators to ration-card-program subsidies in rural areas—as opposed to the responsiveness of the same indicators in urban areas—may be explained by overall lower consumption of

**FIGURE 4.6** Dose-response functions for maternal overweight




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate the probabilities of mothers being overweight, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

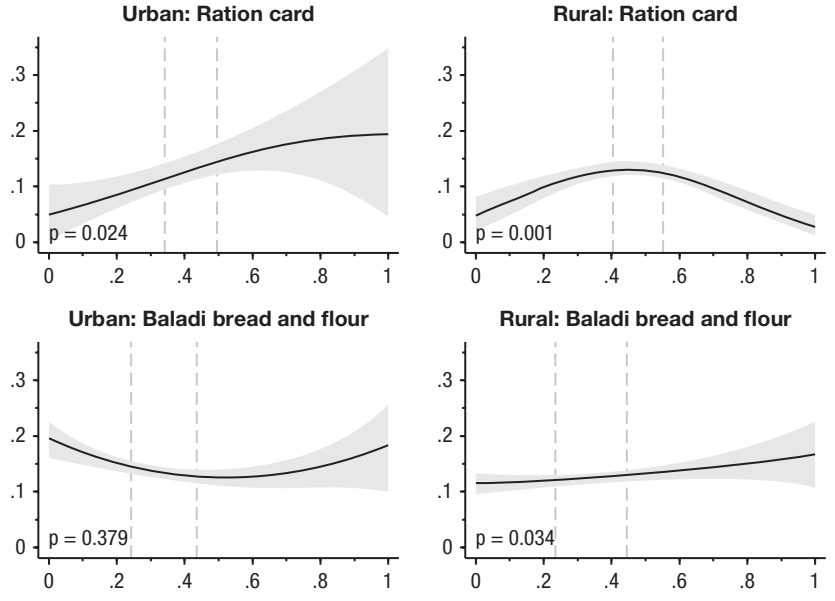
*p* is the p-value of joint significance of the estimated dose-response function (Prob. > chi-sq.). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

subsidized calories because of lower household per capita quotas as well as by higher individual calorie expenditures from physical activity, similar to the case of child overnutrition.

Consistent with the estimation result on child stunting and the estimation result on child overweight, the probability of stunting and overweight coexisting among children increases almost linearly with increasing subsidy levels of the ration card program among urban beneficiary families (at least across most of the subsidy range); among rural beneficiary families, it peaks at lower medium subsidy levels (Figure 4.7). Both dose-response functions have sufficient goodness of fit, and the coefficient estimates that define the shape of most of the curves are statistically significant.[21] Thus, the result for urban beneficiaries provides supportive evidence for our hypothesis that high

21 See Table A.30 in the Appendix.

**FIGURE 4.7** Dose-response functions for child stunting and overweight




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

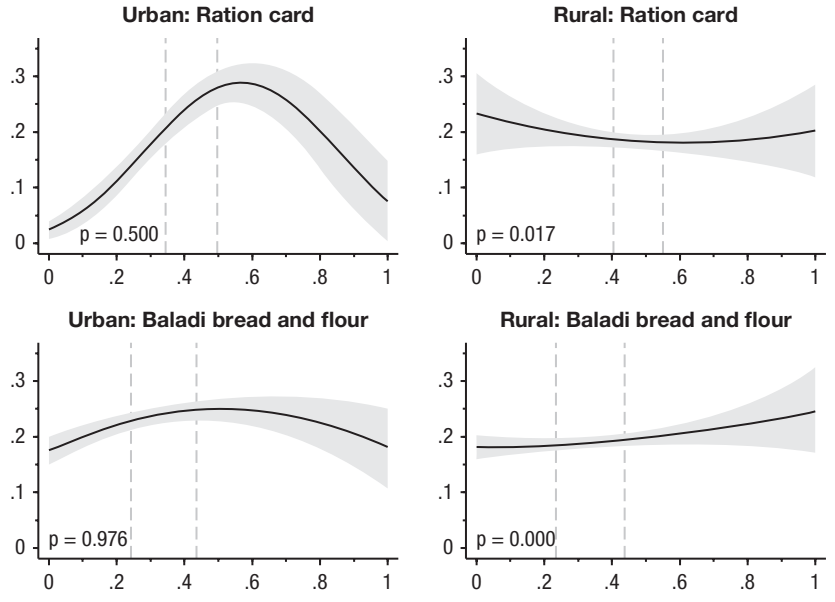
**Note:** The y-axes indicate the probabilities of children being stunted and overweight at the same time, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > chi-sq.). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

ration-card-program subsidies increase the risk of the double burden of malnutrition at the individual level—precisely among young children living in urban areas. Yet the average risk of the double burden of child malnutrition tends to be insignificantly lower among beneficiaries compared to nonbeneficiaries, as the respective ATT estimates above show. Hence, again, it is the received subsidy amounts that matter rather than the ration-card-program participation per se. Although statistically significant, the result for the effect on the double burden of child malnutrition among rural beneficiary families offers little insight, considering that the estimations for the effects on child under- and overnutrition separately produced no clear results.

The dose-response function estimation of the subsidy-level effect of the ration card program on the double burden of malnutrition at the family level among urban beneficiaries yields insufficient goodness of fit that disallows strong interpretation of the estimation result (Figure 4.8). Still, the coefficient

**FIGURE 4.8** Dose-response functions for child stunting and maternal overweight




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate the probabilities of children being stunted and their mothers being overweight, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

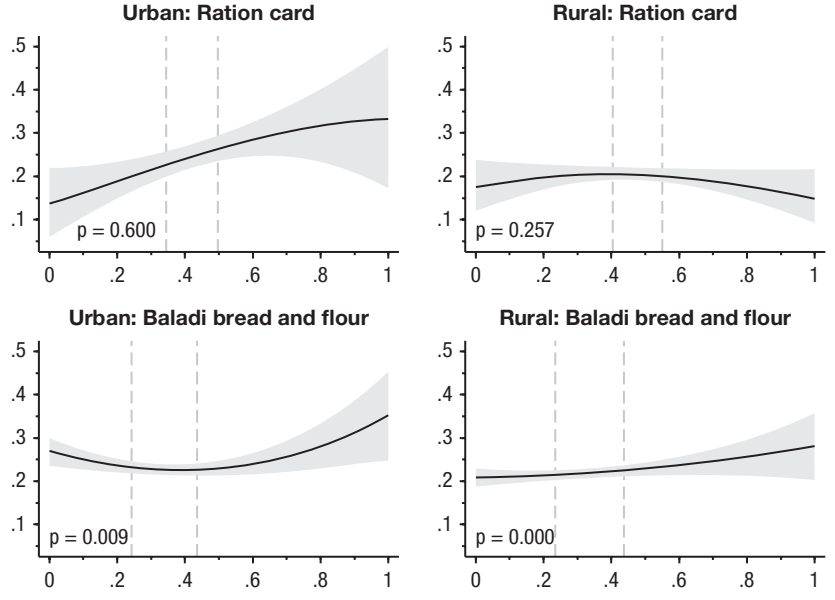
*p* is the p-value of joint significance of the estimated dose-response function (Prob. > chi-sq.). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

estimate of the subsidy-level variable in quadratic form—defining the bell-shaped curve—is statistically significant.[22] The shape of the curve implies that the probability of child stunting and maternal overweight tends to be highest among urban families with medium-high subsidy levels and lower among urban families with low and very high subsidy levels.

The dose-response function for the ration-card-program subsidy-level effect on the double burden of malnutrition at the family level indicates no clear tendency among rural beneficiaries. There is only a slight tendency toward child stunting and maternal overweight being least probable at upper-medium subsidy levels. Although statistically significant, the result is not supported by clear results from the estimations of the subsidy-level effects on child stunting and on maternal overweight as separate indicators, so it should not be over-interpreted. Thus, neither the dose-response model

22 See Table A.31 in the Appendix.

**FIGURE 4.9** Dose-response functions for child overweight and maternal overweight




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate the probabilities of children being overweight and their mothers being overweight, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > chi-sq.). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

estimations for the subsidy-level effects of the ration card program nor the PSM estimations for program participation provide supportive evidence for our hypothesis that the ration card program directly contributes to chronic undernutrition among children and overnutrition among their mothers at the same time.

For the subsidy-level effects on the coexistence of child and maternal overnutrition among urban and rural beneficiary families, the dose-response model estimations yield functions with insufficient goodness-of-fit measures and statistically insignificant coefficient estimates of the subsidy-level variables (Figure 4.9).[23] Nevertheless, the tendency of markedly increasing probability of child and maternal overweight with increasing subsidy levels among urban beneficiary families is consistent with the tendencies found for estimations of the effects on child overweight and maternal overweight as

23 See Table A.32 in the Appendix.

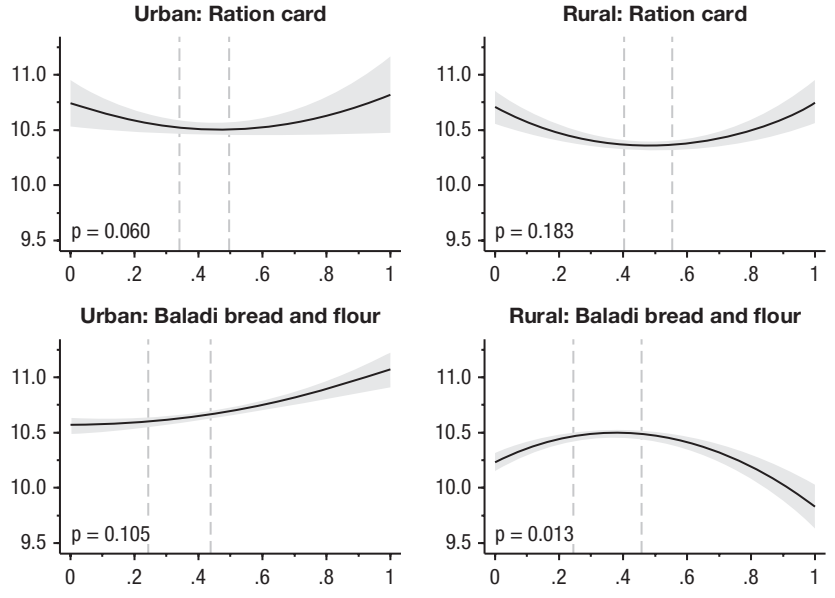
separate indicators. It is also consistent with the result from the corresponding PSM estimation, according to which the average risk of coexisting child and maternal overweight tends to be higher among urban beneficiaries than urban nonbeneficiaries.

Poor household diet quality—both insufficient diversity across all main food groups and infrequent consumption of nutritious food groups—is a likely pathway by which food subsidies can adversely affect child nutrition and hence drive the double burden of malnutrition, as hypothesized. Therefore, the dose-response model estimations for the subsidy-level effects of the ration card program on household diet quality indicators can provide complementary evidence for our hypothesis. The estimation results should be understood in conjunction with the findings from the analysis of Engel curves (presented in the previous chapter), according to which beneficiary and nonbeneficiary families with similar per capita household incomes have similar consumption expenditures for nutritious food groups across the income distribution. However, as noted above, this comparative analysis did not account for different subsidy levels and fell short of demonstrating causality. The dose-response model estimations address these limitations and hence can provide evidence for the hypothesized substitution effects that may occur between subsidized foods and more nutritious, nonsubsidized foods and that increase with rising subsidy levels. This would provide an explanation for the increased incidence of both chronic child undernutrition and child and maternal overnutrition found at high subsidy levels among urban beneficiaries of the ration card program.

The dose-response model estimations for the subsidy-level effects on household dietary diversity among urban and rural beneficiary families do not provide conclusive insights, although the relevant coefficient estimates of the subsidy-level variables are statistically significant (Figure 4.10).[24] Both dose-response functions follow an inverted, flattened bell-shaped curve with the lowest diversity at medium subsidy levels. For urban ration cardholders, the confidence levels of the curve considerably increase toward low and high subsidy levels, so that the curve shape is insufficiently robust for identifying a clear tendency. For the effect on dietary diversity among rural beneficiaries, the estimation yielded insufficient goodness of fit of the dose-response function, restricting its interpretation. It is important to note that the average HDDS at any subsidy level among urban and rural beneficiaries is consistently high, varying only between 10 and 11—out of a maximum of 12. And, the average HDDS among nonbeneficiaries is even slightly higher than

24 See Table A.33 in the Appendix.

**FIGURE 4.10** Dose-response functions for HDDS




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate household dietary diversity scores (HDDSs), and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

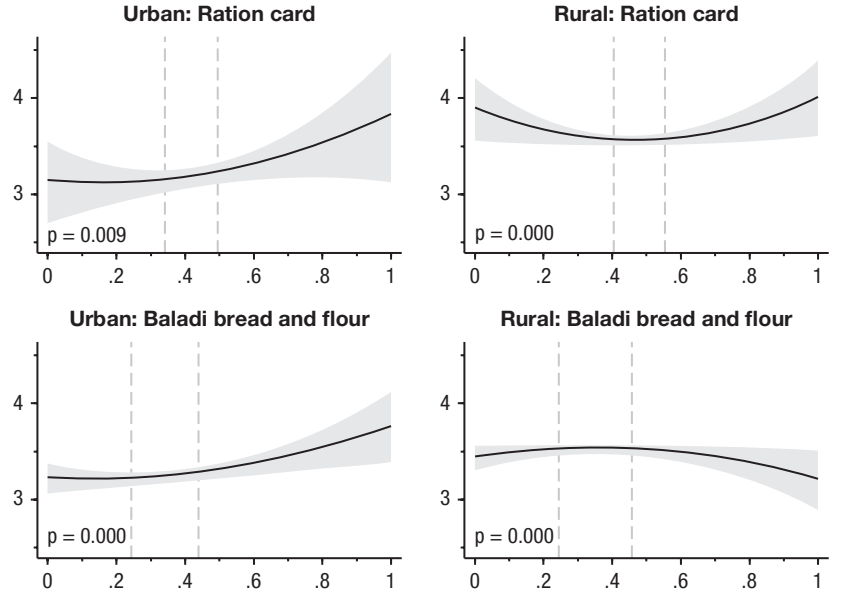
*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

among all beneficiaries, while its variance is quite low overall. Therefore, the HDDS alone appears to be a suboptimal indicator for assessing the effects of the ration card program on household diet quality, given the limitations of our data (Box 4.2).

The dose-response function estimations for the ration-card-program subsidy-level effects on the frequency of household consumption of nutritious food groups yield more revealing and statistically significant results—particularly for urban beneficiary families (Figure 4.11, Figure 4.12, Figure 4.13, and Figure 4.14). The goodness-of-fit measures of all dose-response functions are sufficient, except for legume consumption frequency among rural beneficiary families.[25] The estimation results indicate no clear tendency for the consumption frequency of other nutritious food groups among rural beneficiaries—similar to the results for most nutrition indicators.

25 See Tables A.34–A.37 in the Appendix.

**FIGURE 4.11** Dose-response functions for frequency of household vegetable consumption



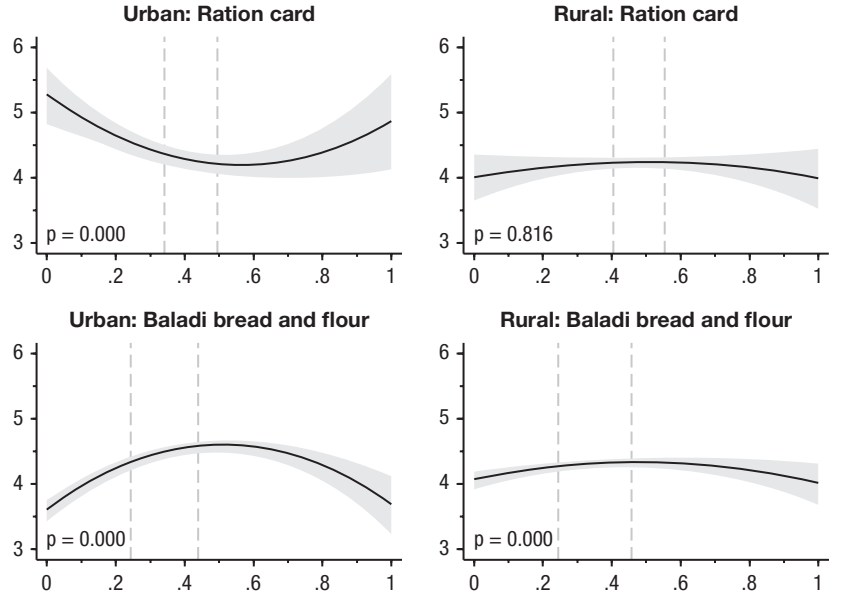

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate frequencies of household vegetable consumption, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

In contrast, the estimation results for urban beneficiaries suggest that, with increasing subsidy levels, they consume vegetables slightly more frequently, but all other micronutrient-rich food groups—including legumes, meat and fish, and milk and dairy products—less frequently (which holds for legume consumption across most of the subsidy range but not for very high subsidy levels). Thus, these findings indeed confirm that urban beneficiaries of the ration card program substitute subsidized calorie-rich and micronutrient-poor foods for more nutritious foods and that this substitution effect increases with increasing subsidy levels, confirming our hypothesis. Animal-source products and legumes are rich in those absorbable micronutrients—especially zinc—and proteins that are important for children's physical and mental development and for preventing child stunting. Therewith, the findings provide an explanation for a likely pathway through which the Egyptian ration card program contributes not only to child and maternal overnutrition but also to chronic child undernutrition in urban areas.

**FIGURE 4.12** Dose-response functions for frequency of household legume consumption



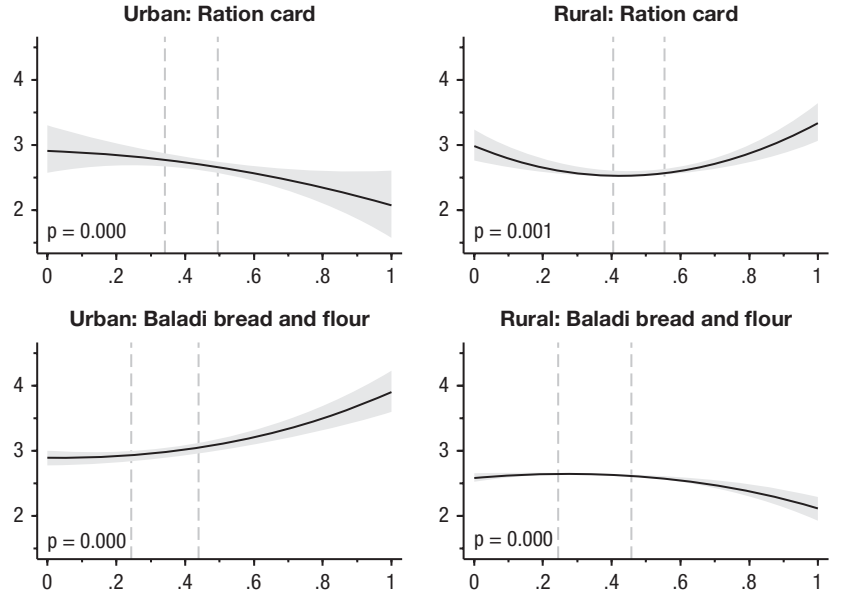

**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate frequencies of household legume consumption, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

Although the frequency of vegetable consumption among urban beneficiaries slightly increases with increasing subsidy levels, it is on average lower among urban beneficiaries than nonbeneficiaries, as the respective ATT estimates above show. The vegetable consumption frequency is generally low—among both beneficiaries and nonbeneficiaries as well as in urban and rural areas. On average, families consume vegetables on less than half of all days during one week, which is less frequently than they consume legumes and milk and dairy products. Such an infrequent vegetable consumption clearly falls short of healthy diet recommendations. Comparing the mean values and variations of the indicators of household dietary diversity and food group consumption frequency suggests that Egyptian diets tend to be inadequate in terms of the daily consumption of micronutrient-rich food groups, especially vegetables (although foods from almost all food groups are consumed over a period of one week, as the HDDS implies). Hence, the dietary shortfall seems to be a matter of low diversity of the regular meals, whereas weekend

**FIGURE 4.13** Dose-response functions for frequency of household meat and fish consumption




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate frequencies of household meat and fish consumption, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

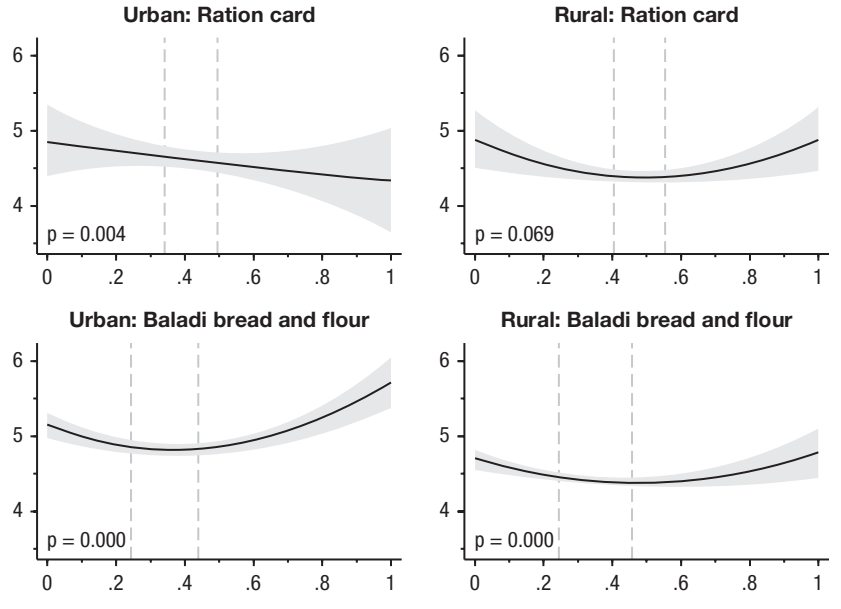
meals may be diverse in different foods, so the HDDS shows high overall dietary diversity.

## Effects of Baladi Bread and Flour Subsidies

The goodness of fit of all GLM estimations of the GPS in the dose-response models for assessing the effects of Baladi bread and flour subsidies is decent (with Pearson scores between 0.20 and 0.26). These measures are higher than those in the dose-response models for assessing the effects of ration-card-program subsidies (possibly because of larger sample sizes).[26] The goodness-of-fit measures of the estimated dose-response functions are similarly low (with R-squared and pseudo R-squared scores of less than 0.13), but

26 Tables A.38–A.41 in the Appendix present the GLM results of the GPS estimates for Baladi bread and flour subsidies.

**FIGURE 4.14** Dose-response functions for frequency of household milk and dairy products consumption




**Source:** Authors' estimation based on data from CAPMAS and WFP (2011).

**Note:** The y-axes indicate frequencies of household milk and dairy products consumption, and the x-axes indicate household subsidy levels of the respective food subsidy programs in the respective residential areas.

*p* is the p-value of joint significance of the estimated dose-response function (Prob. > F). The gray-shaded area marks the 95 percent confidence interval. The vertical dashed lines mark the tertiles of the child samples, separating children from families with low, medium, and high subsidy levels.

they are slightly higher for the dose-response functions of the effects on food group consumption frequency among rural families.[27]

Unlike the ration card program, the Baladi bread and flour program affects child nutrition primarily in rural areas, if at all. The goodness-of-fit measures of all dose-response functions of the effects on child over- or undernutrition in urban areas are insufficient, and all coefficient estimates of the subsidy-level variables are statistically insignificant, disallowing strong interpretation of the results. In addition, they do not indicate any clear tendency for child nutrition in urban areas.

The dose-response function of the subsidy-level effect on child HAZ in rural areas indicates that a child's body height (relative to age) tends to

27 Tables A.42–A.55 in the Appendix present the dose-response function estimates for the effects of Baladi bread and flour subsidies.

decrease with increasing Baladi bread and flour subsidy levels at slightly declining margins (Figure 4.1). However, the probability of child stunting in rural areas is quite unresponsive to different subsidy levels, as the respective dose-response functions show (Figure 4.2). Both dose-response functions have statistically significant coefficient estimates of the subsidy-level variables, but the goodness of fit of the function for the effect on child HAZ is insufficient.[28]

The Baladi bread and flour program also seems to contribute to child overnutrition in rural areas. Child BMIZ tends to linearly increase with increasing subsidy levels (Figure 4.3), and the probability of child overnutrition tends to rise especially among families with high subsidy levels (Figure 4.4). However, it should be noted that the coefficient estimates of the subsidy-level variables in both dose-response functions are statistically insignificant (at common levels), and the goodness of fit of the function for the effect on child overweight is insufficient, too.[29]

Like the ration card program, the Baladi bread and flour program affects maternal overnutrition among urban families, whereas mothers' BMIs and the probability of maternal overweight among rural families are unresponsive to different levels of the Baladi bread and flour subsidy (Figure 4.5 and Figure 4.6).[30] The dose-response functions indicate that both mothers' BMIs and the probability of maternal overweight in urban areas tend to (almost) linearly increase with increasing subsidy levels. Thus, our estimation results consistently show that maternal overnutrition among urban families in Egypt is partly caused by food subsidies from both subsidy programs, while the effect of the ration card program on maternal overweight appears to be larger than that of the Baladi bread and flour program.

The double burden of malnutrition at the individual and the family levels in rural areas is largely unresponsive to different subsidy levels of the Baladi bread and flour program (Figure 4.7 and Figure 4.8). This also holds true for the incidence of child and maternal overweight in the same family in rural areas (Figure 4.9). These results are robust and consistent with the estimation results for child stunting and child and maternal overweight among rural families. In contrast, low goodness-of-fit measures and statistically insignificant coefficient estimates of the subsidy-level variables do not permit any interpretation of the dose-response functions of the subsidy-level effects on the double

28 See Tables A.42 and A.43 in the Appendix.

29 See Tables A.44 and A.45 in the Appendix.

30 See Tables A.46 and A.47 in the Appendix.

burden of malnutrition at the individual and the family levels in urban areas.[31] Thus, we do not find evidence to support our hypothesis that Baladi bread and flour subsidies contribute directly to the double burden of malnutrition. The dose-response function of the subsidy-level effect of the Baladi bread and flour program on the probability of coexisting child and maternal overweight in urban areas indicates a slightly decreasing probability up to a medium subsidy level and then a slightly increasing probability at higher subsidy levels.[32] However, this result is at best only weakly supported by the estimation results for child overweight and maternal overweight among urban families.

Keeping the aforementioned limitations of the HDDS in the context of this analysis in mind, we see that the dose-response function of the subsidy-level effect of the Baladi bread and flour program on household dietary diversity indicates a notable and statistically significant tendency for rural families.[33] It indicates that dietary diversity among rural families first slightly increases up to a medium subsidy level and decreases with higher subsidy levels at higher margins (Figure 4.10). According to that, there is an optimal, medium level for Baladi bread and flour subsidies in terms of their beneficial effect on household dietary diversity, while low subsidies have a less adverse dietary effect than high subsidies. Hence, above a certain subsidy level, higher amounts of cheap bread and flour tend to crowd out other—perhaps more nutritious—food groups and thereby likely contribute to overall unbalanced diets among rural families. The dose-response function of the subsidy-level effect on household dietary diversity among urban families indicates a tendency toward increasing dietary diversity with increasing subsidy levels. This tendency is not robust, however, given insignificant coefficient estimates of the subsidy-level variables. It should be noted that rural families have a significantly lower HDDS than that of urban families, and the indicator's variance is larger among rural families, which may help finding statistically significant results.

The dose-response model estimations for the subsidy-level effects of the Baladi bread and flour program on household consumption frequency of nutritious food groups yields ambiguous results. The frequencies of vegetable consumption and meat and fish consumption tend to slightly increase among urban families and slightly decrease among rural families at high (or medium-high) subsidy levels (Figure 4.11 and Figure 4.13). Yet the coefficient

31 See Tables A.48 and A.49 in the Appendix.

32 See Table A.50 in the Appendix.

33 See Table A.51 in the Appendix.

estimates of the subsidy-level variables are statistically insignificant in the dose-response functions for the effects among urban families.[34] At low and medium subsidy levels, the Baladi bread and flour program contributes to more frequent legume consumption among urban families; at high subsidy levels, however, the program contributes to less frequent legume consumption among urban families (Figure 4.12). The frequency of legume consumption among rural families seems to be unresponsive to different subsidy levels. Among both urban and rural families, the frequency of milk and dairy products consumption slightly decreases with increasing Baladi bread and flour subsidy levels across most of the subsidy range and increases only at very high subsidy levels again (Figure 4.14).

Overall, we found few consistent responses in the nutrition and diet quality indicators to different subsidy levels of the Baladi bread and flour program—with the important exceptions of child overnutrition among rural families and maternal overnutrition among urban families. Given that we did not find clear evidence for the effects of Baladi bread and flour subsidies on chronic child undernutrition, the dose-response estimations for the effects on household diet quality indicators have limited explanatory potential for the possible pathway through which these food subsidies could affect that nutrition outcome. In this sense, the ambiguous results of the dose-response model estimations for the effects on household diet quality are consistent with those for the effects on child stunting and the double burden of malnutrition at the individual and the family levels. A possible explanation for the disparity in the responsiveness of nutrition and diet quality indicators to the subsidies from the ration card program (among urban families) and the Baladi bread and flour program may relate to the fundamental differences in the design of the two programs (discussed in the previous chapter). Probably most crucially, Baladi bread is indeed an inferior good—unlike subsidized rice, sugar, and cooking oil. Further, the lack of (self-)targeting of the ration-card-program subsidies to the people in need and the selection of subsidized foods misaligned with people's real nutritional needs are what drive adverse nutritional outcomes.

34 See Tables A.51–A.55 in the Appendix.