Economic analysis of the link between diet quality and health: Evidence from Kosovar micro-data

Kushtrim Braha[1], Andrej Cupák[2], Artan Qineti[1], Ján Pokrivčák[1], Marian Rizov[3]

[1] Slovak University of Agriculture, Slovakia

[2] National Bank of Slovakia, Slovakia

[3] University of Lincoln, United Kingdom

Contribution presented at the XV EAAE Congress, "Towards Sustainable Agri-food Systems: Balancing Between Markets and Society"

August 29th – September 1st, 2017

Parma, Italy

Economic analysis of the link between diet quality and health: Evidence from Kosovar micro-data*

Abstract

We analyze the link between diet quality and health outcomes measured by body-mass index (BMI) in a sample of 8,900 Kosovar individuals utilizing household expenditure micro-data. Using a household model of health production we devise a two-stage empirical strategy to estimate the antecedents of diet diversity and its effect on BMI. Economic factors and demographic characteristics play an important role in the choice of more balanced diets. Results from the BMI analysis support the hypothesis that diet diversity is associated with optimal BMI and thus healthier status. One standard deviation increase in diet diversity leads to 2.3% increase in BMI of the underweight individuals and to 1.4% reduction in BMI of the obese individuals. The findings have important implications for food policies aiming at enhancing the public health in Kosovo.

Keywords: BMI, diet diversity, diet quality, health, Kosovo

1 Introduction

Despite progress made in recent years a significant number of people suffer from food insecurity and undernourishment (FAO, 2015). Countries where food expenditures constitute significant share of households' incomes are the most vulnerable (Clapp and Cohen, 2009; McMichael and Schneider, 2011). Food insecurity is closely associated with suboptimal nutrition, and consequently worsening health status (Hatloy et al., 1998; Thomas and Frankenberg, 2002).

Darmon and Drewnowski (2008) note that energy-dense and nutrient-poor diets are cheap and more common for those with limited means; socio-economic conditions are therefore important predictors of diet quality. Diets of many households in transition countries are particularly monotonous, mainly based on cheap cereal products (Swinnen and Van Herck, 2011). Quality of diet in transition countries depends heavily on real incomes, prices, and household socio-economic characteristics (e.g. Moon et al., 2002; Herzfeld et al., 2014; Cupák et al., 2016).

Figure 1 depicts the link between economic development and health suggesting possible non-linearity due to suboptimal nutrition leading, in turn to rise in obesity and deteriorating health throughout the development process as countries become richer. Overweight or obesity is more

* We thank the Kosovar Statistical Office for granting access to the Household Budget Survey data. The authors are solely responsible for the content of the paper. The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the National Bank of Slovakia.

widespread in the middle-income Central and Eastern European (CEE) countries compared to the high-income, more developed European countries. Rising trends in overweight and obesity in many of the Eastern European countries have been also shown in a descriptive study by Knai et al. (2007) arguing that up to 6 % of the total healthcare costs can be attributed directly or indirectly to obesity and consequent illnesses which can have important implications for the population's economic productivity.

[Figure 1 about here]

Whereas the individual health status, measured by the body-mass index (BMI), and its determinants have been studied in a number of European countries with detailed micro-data (see, for example, Gutiérrez-Fisac, 1999 for Spain; Ali and Lindström, 2005 for Sweden; or Kleiser et al., 2009 for Germany), the list of studies for the transition CEE countries is very limited. Some exceptions are studies based on Russian individual data evaluating food demand patterns and rise of obesity in Russia (e.g. Huffman and Rizov, 2007; 2010; Herzfeld et al., 2014).[1] This presents a significant gap in the literature as economic development of transition countries has not reached the Western European levels yet which can have important implications for the food security status and health of households and individuals.

Kosovo offers an important case study of the link between food security, diet, and health from both scientific and policy intervention viewpoint. Kosovo is one of the poorest transition CEE countries, with per capita GDP at about 3,000 Euro; 29.7% of its population lives below the poverty line and 10.2% of the population lives in extreme poverty (World Bank/KAS, 2011). Although a significant GDP growth of 4.5 percent annually is marked since the early 2000s, growth rates have had limited impact on poverty reduction and public health improvement (UNDP 2014; World Bank, 2015). Thus, the food security situation in Kosovo and its impact on health remains an important development issue. However, so far the focus in the case of Kosovo has mostly been on the contribution of agriculture to economic development and food availability as a part of the food security agenda (see Osmani et al., 2013; MAFRD, 2014; Braha et al., 2015).

The ultimate goal of this paper is to analyze the determinants of individual health status, measured by the BMI, with the main focus on diet quality, proxied by different diet diversity measures. Our analysis is based on the individual Kosovar micro-data from 2012. The nutrition literature (e.g. Hatloy et al., 1998; Darmon and Drewnowski, 2008) shows that consumption of diverse diet has positive impact on health and the diet diversity is a good indicator of food security

[1] There are no relevant studies on the link between food consumption, BMI and health in developing country context. An exception is a recent paper utilizing household budget survey data to analyze the impact of food expenditure patterns on the child anthropometric measurements in four developing countries (Humphries et al., 2017).

for a particular household. We estimate demand for diet diversity and its impact on individual health measured by obesity status (BMI) by means of OLS and quantile regressions. This is, to the best of our knowledge, the first attempt to evaluate diet quality and test the link between diet and health in CEE countries using representative micro-data. We use an innovative empirical framework based on a sound theoretical model of health production following Huffman and Rizov (2010). We find that diverse diet positively affects population's health in the lower quantiles of the BMI's distribution, it has no effect in the median of the distribution, and has significant negative effect in the upper quantiles of the distribution. Other control variables like gender, age, education, and propensity to smoke also significantly influence the individual health status. The finding of an inverted-U shape relationship between the diet quality and BMI can have important implications for policies designed to prevent undernourishment and enhance healthy lifestyles of individuals.

The paper is organized as follows. In section 2 we present the theoretical framework underlying the empirical analysis in section 3. Section 4 reports and discusses estimation results while section 5 concludes the paper.

2 Theoretical framework

The productive household models of health developed by Rosenzweig and Schultz (1982) and Grossman (2000) provide a useful theoretical background. Huffman and Rizov (2010) building on the existing theory develop an empirical framework that explicitly models the link between food diet and health outcomes as measured by BMI. An important proposition in the framework is that the health status of each household member is represented and determined by the degree of his/her overweight or obesity. A monotonic negative relationship between the obesity measure and health over a relevant range of the distribution of obesity is assumed. The assumption is quite reasonable for populations of medium to high income countries experiencing widespread obesity problems. In low to medium income countries such as Kosovo the relationships between diet quality, measures of overweight and obesity, and health might be non-linear, with inverted-U shape, a relationship that we explicitly account for in the estimation stage.

We start by specifying an individual utility function

$$U = U(D, C, BMI, L; O). \quad (1)$$

Utility is determined by the food diet consumed, *D*; consumption of other goods (excluding food) and services, *C*; body-mass index, *BMI* proxying for health status; leisure, *L*; and fixed characteristics, such as age, gender, education, and socio-economic background, *O*.

The individual has a BMI production function

$$BMI = B(D, L, O, \varepsilon), \quad (2)$$

where ε is the unobserved individual characteristics that affect the individual's BMI; such characteristics may include genetic factors. In large meta-population samples it is likely that ε is randomly distributed, with a zero mean, and influences both younger and older population cohorts in a similar manner (see for example, Malis et al., 2005; Dolton and Xiao, 2017). Food consumption affects utility directly and indirectly, through BMI production, by providing energy, vitamins and minerals.

The individual has a budget constraint

$$P_D D + P_C C = W(T - L) + N\,, \tag{3}$$

where P_D and P_C denote the prices of food (*D*), and other goods and services (*C*), respectively; *W* is the wage rate per unit of time, *T* is the fixed time endowment (*T*-*L*=work), and *N* is the non-labor income.

For an interior solution of the model, we substitute equation (2) into (1) and use the budget constraint (3). The individual chooses *D*, *L* and *C* by maximizing his/her utility subject to the budget constraint. The utility maximization problem can be written as

$$\Lambda = U[D, C, B(D, L; O, \varepsilon), L; O] + \lambda(WT + N - P_D D - P_C C - WL)\,, \tag{4}$$

where λ is the Lagrange multiplier representing the marginal utility of individual's full income. The first order conditions for an optimal solution are:

$$U_B B_D + U_D = \lambda P_D\,, \tag{5}$$

$$U_B B_L + U_L = \lambda W\,, \tag{6}$$

$$U_C = \lambda P_C\,, \tag{7}$$

$$WT + N = P_D D + P_C C + WL\,, \tag{8}$$

where $U_B = \partial U / \partial B,\ B_D = \partial B / \partial D,\ U_D = \partial U / \partial D,\ \ B_L = \partial B / \partial L,\ \ U_L = \partial U / \partial L$ and $U_C = \partial U / \partial C$. For an interior solution, equations (5)-(8) yield the individual's optimal demand functions for *D*, *L* and *C*:

$$\Phi^* = f_\Phi(P_D, P_C, W, N, O, \varepsilon_\Phi), \qquad \Phi = D, L, C. \tag{9}$$

Therefore, the demand for inputs into the BMI production function depends on the prices of the purchased inputs (P_D, P_C), the wage rate (*W*), non-labor income (*N*), fixed factors (*O*) and unobserved factors (ε), which are assumed to have zero expected mean. After substituting the optimal demand functions D^* and L^* from equation (9) into the BMI production function (2), we obtain the individual's BMI supply function:

$$BMI^* = B_S(P_D, P_C, W, N, O, \varepsilon_B)\,. \tag{10}$$

Note that the BMI supply function (equation 10) is a reduced-form (behavioral) relationship based on the optimal individual's decisions while the individual's BMI production function (equation 2) is a technology relationship. Equation (10) represents the solution to the first-order (Kuhn-Tucker) conditions for the structural endogenous variables (*D*, *L*, *C*) in terms of the exogenous factors which include wages, prices, and characteristics of the BMI production and utility functions. This is a common approach of transitioning to an empirical framework.

An alternative (structural) approach to the transition to an empirical framework that is particularly suitable for our goal to estimate the impact of diet quality on the health status can be implemented in two stages following Huffman and Rizov (2010). First, we estimate demand function equations (9) for diet diversity (and smoking as an important factor affecting both diet and health) and a wage equation to obtain a proxy of leisure demand. Second, we substitute the predicted values from the first stage in the technology equation (2). Given issues with data availability, specifically the lack of direct price information, we adopt the structural approach in this paper. Thus, in the empirical analysis, we first estimate an optimal diet diversity (demand) equation following Herzfeld et al. (2014). Next, in a second step, using the predicted diet diversity indicator, together with exogenous controls for the individual and household behavior listed in vector *O*, we estimate the BMI supply function:

$$BMI^{*} = B_{S}^{*}(D^{*}, L^{*}, O, \varepsilon_{B}^{'}) \qquad (11)$$

Equation (11) represents the link between diet quality and health status and is in the focus of our empirical analysis that follows. Additional advantage of the two stage procedure adopted is that it deals well with the possible endogeneity of explanatory variables used in the BMI supply equation.[2] This is achieved by using predicted values from the first stage where as instruments are added appropriate exogenous variables following the relevant literature.

3 Data and estimation strategy

We analyze food diet quality of Kosovar households and its impact on individuals' health status using the Household Budget Survey (HBS) data collected by the Kosovar Statistical Office. Our dataset consists of four seasonal rounds in 2012. The survey provides detailed information on household incomes and expenditures on food and non-food goods and services. The HBS data also contains detailed information on quantities consumed by each household, location of the household,

[2] In a recent paper Morales et al. (2016) address comprehensively in a simultaneous equation framework the endogeneity issues of estimating the determinants of obesity using rich U.S. panel data as they focus on the endogenous lifestyle and location choices. Given the limitations of our data and focus of our analysis on both the determinants of diet diversity and, in turn, the impact of diet on health status proxied by BMI, we consider our two-stage approach most appropriate.

and its size. Individual household member characteristics such as age, education, and working status are also available. The 2012 total sample contains approximately 8,900 adult individuals. The information on food consumption is collected on a one-month recall basis in four waves, one for each of the four seasons. Importantly, data also contains anthropometric information on individual weight and height, and lifestyle patterns such as frequency and nature of physical activity, and tobacco smoking.

3.1 Variables

There is a plethora of definitions and measurements of diet quality in the literature. Hoddinott (1999) offers a comprehensive overview of frameworks for measuring household diet quality and food security. In this paper we employ three diversity measures of household diet quality: (i) the count measure of food items (CM), (ii) diversity measured by the Simpson index (SI), and (iii) diversity measured by the Entropy index (EI).

The number of food items consumed during specific time period has been commonly used as an indicator of diet diversity (e.g., Jackson 1984; Kant 1996); it can be defined as $CM = \sum_{j=0}^{N} Q_i$, where CM (count measure) is a sum of dummy variables Q_i taking value of 1 if a household consumes *i*-th food item in its food basket during the survey recall period and 0 otherwise. The other two measures, which have also become popular in measuring diet diversity in the food economics literature (e.g. Theil and Finke, 1983; Thiele and Weiss, 2003; Hertzfeld et al., 2014) are the Simpson index defined as $SI = 1 - \sum w_i^2$ and the Entropy index defined as $EI = \sum w_i \log(1/w_i)$, where w_i is the budget share of the *i*-th (disaggregated) food item in the total food expenditure (basket). Unlike the count measure of diet diversity, Simpson and Entropy indices take into account the distribution of food consumption. The formulation of SI and EI implies that diversity is higher when more food items are consumed in equal (quantity or expenditure) proportions such that a higher value of the index indicates a more diverse diet.

The second stage of our analysis focuses on the impact of diet quality on individual health status in Kosovo. We compute an anthropometric measure - body mass index (BMI) - by dividing individual weight (in kilograms) by the square of individual height (in meters).[3] Anthropometric studies suggest measurement of the physical dimensions and gross composition of the human body as the most powerful tools to determine the long-term individual nutrition and health status (Victora, 1992; Gibson, 2005; Neufeld and Osendarp, 2014). An added value of such indicators lays

[3] According to WHO (2014) the optimal BMI ranges from 18.5 to 25, where BMI lower than 18.5 indicates underweight, while higher BMI score than 25 indicates overweight; BMI above 30 indicates obesity status.

in their ability to discriminate between different physiological and biological factors (Gorstein et al., 1994). Poor nutritional status influences health and wellbeing through the life cycle from the prenatal period on into elder years (Cook and Frank, 2008). Therefore, anthropometric indicators are determined as impact indicators that assign the level to which the food is utilized and converted into satisfactory nutrition status (Reinhard and Wijayaratne, 2000). Figure 2 presents the BMI index distribution of Kosovar adults by gender.

[Figure 2 about here]

Descriptive statistics of the dependent variables discussed above and all explanatory variables as discussed in the theoretical section (components of vector *O*) and, in addition, regional dummies, proxing for prices, used in stage one are reported in Tables A.1, A.2, and A.3. Table 1 reports descriptive statistics of variables entering the stage two analysis based on the BMI supply equation.

[Table 1 about here]

3.2 Estimation strategy

Following our theoretical model, we estimate in a first stage the endogenous demand variables – diet diversity, smoking, and leisure – affecting BMI production and supply and then use their predicted values in a second stage BMI supply equations.

Diet diversity

Theoretically, the diet diversity specification is based on standard demand analysis and extensions by Jackson (1984), Stewart and Harris (2005), and Herzfeld et al. (2014). We empirically implement the household diet diversity demand function following Herzfeld et al. (2014) and Cupák et al. (2016) by specifying an estimating equation where household diet diversity is explained by prices, controlled for by regional dummy variables, household income, and household characteristics (household size, composition, education level, and employment status), and production and consumption patterns in terms of household consuming own food production and/or food away from home. As controls we also add season and region dummy variable sets. Detailed results of estimating OLS models with each of the three measures of diet diversity (CM, SI, EI) as dependent variable are reported in an Appendix (Table A.4).

Propensity of smoking

We estimate propensity of smoking as a function of prices (regional dummy variables) and income as well as of individual characteristics listed in vector *O*. The literature on myopic addiction initiated by Pollack (1970) models smoking as a partial adjustment model where a lagged dependent variable represents the propensity of smoking which is carried over from period to period and its coefficient can be interpreted as an indicator of the strength of addiction. Because of the cross-

sectional nature of our analysis and issues with availability of data we include a categorical variable capturing physical activity and lifestyle patterns of individuals, thus, arguably, also capturing the variation in the degree of addiction across smokers. Furthermore, the rational addiction model of Becker and Murphy (1988) implies that the actual value (propensity) of future smoking should be included in the regression as well. Due to lack of appropriate information in our data we are not able to estimate the model of propensity of smoking by fully controlling for rational addiction behavior. Becker et al. (1991), however, suggest that the long-run responses obtained from both myopic and rational addiction models are similar. Detailed results of estimating Probit model are reported in an Appendix (Table A.5).

Leisure demand

As an approximation of leisure demand we estimate a wage equation. Considering that our main goal is to analyze the link between diet quality and BMI, an estimate of opportunity cost of time given that leisure is a normal good is a reasonable control for leisure demand. Wage equation is specified following Becker (1965) and is estimated following Heckman (1974). The predicted wage rate for all sampled individuals is used to control for leisure demand in the second-stage BMI supply equation. The dependent variable in the wage equation is log of the wage rate and the explanatory variables are individual and household characteristics (as specified in vector *O*) plus seasonal and regional dummies. Number of adult household members, number of children in the household, and non-labor income, including remittances, and controls for constraints and incentives of an individual to undertake market employment are used as identifying variables, in the first step. Regional fixed effects control for relative labor market conditions and prices of food and other omitted variables that differ by region. The wage equation is estimated using the Heckman selection model. Full estimation results are reported in Appendix - Table A.6.

BMI supply

In the second stage of our empirical analysis, we use the predicted values of diet diversity, propensity of smoking, and wage rate to estimate the BMI supply function. The following are also included as explanatory variables: individual characteristics as specified in vector *O*, set of dummy variables controlling for type of physical activity and lifestyle, and dummies indicating urban areas and capital city location.[4] The BMI supply equation (11) is estimated by standard OLS regression. Considering our theoretical discussion and expectations about a non-linear, inverted-U shape link

[4] The importance of location for the overweight and obesity status has been explicitly recognised in recent studies by Morales et al. (2016) and Raftopoulou (2017) as the former paper models the location decision as endogenous choice. In our cross-sectional analysis regional dummies are exogenous controls as well as the physical activity variable is which we consider as representing medium-term (at least) lifestyle patterns associated with individuals' working lives given the specific survey question in our data.

between diet diversity and overweight and obesity, we also estimate quantile regressions (QR) which form the core of our estimation strategy.[5]

Given the cross-sectional nature of our data, to address concerns with endogeneity problems when estimating BMI supply equations we introduce the explanatory variables stepwise, one-by-one, to check for stability of coefficients to changes in the specification. The procedure yields robust coefficients and gives us the confidence to conclude that our estimation strategy deals reasonably well with endogeneity problems. Furthermore, in equilibrium, it is reasonable to assume that factors affecting obesity are predetermined, that is, even though obesity affects an individual's characteristics, an individual's characteristics (and other behavioral and environmental factors) determine obesity. Important in this relationship are the lags of the effects. We argue here that the time lag of the obesity effect on an individual's characteristics is much longer than the lag of individual characteristics' effects on obesity.

4 The link between diet diversity and BMI

In Tables 2, 3, 4 are reported the main estimation results from quantile regressions for each of the three (predicted) diet diversity measures. We focus on three quantiles of the BMI distribution (Q=0.10, Q=0.50, and Q=0.90) representing the three main categories of individual's health status– underweight, normal weight, and obese. The explanatory variable of main interest in each specification is the predicted diet diversity measure. The most important message from our estimation results is that for all three diet diversity measures there is a non-linear, inverted-U shape relationship between diet diversity and BMI.[6] In the underweight quantile the relationship is positive and statistically significant suggesting that better (more diverse) diet is associated with higher BMI which indicates a better health in this range of the BMI distribution. In the median quantile representing individuals with normal weight the impact of diet on BMI is positive but not statistically significant. In the obese quantile of the BMI distribution, where the risk or diabetes and cardio vascular despises is higher, we find statistically significant negative relationship between diet diversity and BMI indicating again that better diet leads to better health by supporting optimal weight. In terms of measured impact, considering the whole original distributions, one standard

[5] In a different but related context Dolton and Xiao (2017) study the intergenerational transmission of BMI within a quantile regression framework and find substantial differences across the distribution.

[6] The quantile differences in the diet diversity coefficients are statistically significant at 5% level or better as confirmed by Wald test results which are available on request.

deviation increase in diet diversity leads to between 2.3% and 2.4% increase in BMI of the underweight individuals and between 1.2% and 1.6% reduction in BMI of the obese individuals.[7]

[Tables 2, 3, 4 about here]

The rest of the estimation results associated with individual demographic characteristics are in line with existing theories and empirical evidence. While (general) education appears to have overall negative effect on BMI in the median quantile, in the severely obese category it has no statistically significant effect. The demand for leisure does not appear to significantly impact on BMI while smoking has significant negative association with BMI in all parts of the distribution. Physical activity has a statistically significant positive association with BMI only for the median part of the distribution as individuals involved actively in sports are characterized by increased muscle mass and BMI. Urban areas seem to contain more individuals from both extremes of the BMI distribution but when the capital city region is considered the individuals residing there seem to have more balanced (less dispersed) BMI distribution.

5 Conclusion

Food insecurity in Kosovo is revealed by the significant dispersion of measures of dietary quality computed from household expenditure micro-data as well as by the dispersion of anthropometric indicators such as BMI. Building on a household model of health production we devise a two-stage estimation strategy to analyze the antecedents of balanced diet and its impact on BMI indicating individual health status. From our first-stage estimation results concerning diet diversity we can conclude that Kosovars are significantly exposed to food security risks caused mainly by low incomes, high rates of poverty, low education, and predominantly living in the rural areas.

Our second-stage BMI supply estimation results suggest that diet diversity as an indicator of diet quality has an important role to play in tackling underweight and obesity problems both associated with food and nutritional insecurity. Given that underweight and obesity problems signal ill health hazard, policies facilitating better quality, more diverse, diet and healthy lifestyles may have important positive impact on public health in Kosovo.

Acknowledgement

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-15-0552 and No. APVV-16-0321. The authors also acknowledge financial support from the projects VEGA 1/0806/15, VEGA 1/0930/15, VEGA 1/0797/16, VEGA 1/0935/17, VEGA 1/0928/17.

[7] The standard deviations of the diet diversity distributions by the three quantiles used are very similar for each diet diversity measure; these results are available from the authors.

References

Ali, S.M. and Lindström, M. (2006). Socioeconomic, psychosocial, behavioural, and psychological determinants of BMI among young women: differing patterns for underweight and overweight/obesity. *The European Journal of Public Health*, 16 (3): 324–330.

Becker, G.S. (1965). A theory of the allocation of time. *The Economic Journal*, 75 (299): 493–517.

Becker, G.S. and Murphy, K.M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96 (4): 675–700.

Becker, G.S., Grossman, M. and Murphy, K.M. (1991). Rational addiction and the effect of price on consumption. *American Economic Review*, 81 (2): 237–241.

Braha, K., Rajcaniova, M. and Qineti, A. (2015). Spatial Price Transmission and Food Security: The case of Kosovo. Paper presented in AIEAA (Italian Association of Agricultural and Applied Economics) Fourth Congress, June 11-12, Ancona, Italy.

Clapp, J. and Cohen, M. (2009). *The global food crisis: Governance, challenges and opportunities*. Waterloo: Wilfrid Laurier University Press.

Cook, J. T. and Frank, D. A. (2008). Food security, poverty, and human development in the United States. *Annals of the New York Academy of Sciences*, 1136 (1): 193–209.

Cupák, A., Pokrivčák, J. and Rizov, M. (2016). Diversity of food consumption in Slovakia. *Politická ekonomie*, 64 (5): 608–626.

Darmon, N. and Drewnowski. A. (2008). Does social class predict diet quality? *The American Journal of Clinical Nutrition*, 87 (5): 1107–1117.

Dolton, P. and Xiao, M. (2017). The intergenerational transmission of body mass index across countries. *Economics & Human Biology*, 24, 140–152.

FAO (2015). The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome: FAO.

Gibson, R.S. (2005). *Principles of nutritional assessment*. Oxford university press, USA.

Gorstein, J., Sullivan, K., Yip, R., De Onis, M., Trowbridge, F., Fajans, P. and Clugston, G. (1994). Issues in the assessment of nutritional status using anthropometry. *Bulletin of the World Health Organization*, 72 (2): 273–283.

Grossman, M. (2000). The Human Capital Model. In A.J. Cutler and J.P. Newhouse, Handbook of Health Economics, Vol. 1A, New York, NY: Elsevier, pp. 248–408.

Gutiérrez-Fisac, J.L., Rodríguez Artalejo, F., Guallar-Castillon, P., Banegas Banegas, J. R. and del Rey Calero, J. (1999). Determinants of geographical variations in body mass index (BMI) and obesity in Spain. *International Journal of Obesity*, 23 (4): 342–347.

Hatloy, A., Torheim, L. E. and Oshaug, A. (1998). Food variety: a good indicator of nutritional adequacy of the diet. A case study from an urban area in Mali, West Africa. *European Journal of Clinical Nutrition*, 52 (12): 891–898.

Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, 42 (4): 679–694.

Herzfeld, T., Huffman, S. and Rizov, M. (2014). The dynamics of food, alcohol and cigarette consumption in Russia during transition. *Economics & Human Biology*, 13 (March): 128–143.

Hoddinott, J. (1999). Choosing outcome indicators of household food security, International Food Policy Research Institute (IFPRI), Washington D.C.

Huffman, S. and Rizov M. (2007). Determinants of obesity in transition economies: The case of Russia. *Economics & Human Biology*, 5 (3): 379-391.

Huffman, S. and Rizov, M. (2010). The rise of obesity in transition: Theory and empirical evidence from Russia. *Journal of Development Studies*, 46 (3): 574–594.

Humphriesa, D.L., Deardenb, K.A., Crookstonc, B.T., Woldehannad, T., Pennye, M.E., Behrmanf, J.R. (2017). Household food group expenditure patterns are associated with child anthropometry at ages 5, 8 and 12 years in Ethiopia, India, Peru and Vietnam. *Economics & Human Biology*, 26: 30–41.

Jackson, L. (1984). Hierarchic demand and the Engel curve for variety. *The Review of Economics and Statistics*, 66 (1): 8–15.

Kant, A.K. (1996). Indexes of overall diet quality: a review. *Journal of American Dietetic Association*, 96 (8): 785-791.

Kleiser, C., Rosario, A. S., Mensink, G. B., Prinz-Langenohl, R. and Kurth, B. M. (2009). Potential determinants of obesity among children and adolescents in Germany: results from the cross-sectional KiGGS Study. *BMC Public Health*, 9: 46.

Knai, C., Suhrcke, M., and Lobstein, T. (2007). Obesity in Eastern Europe: an overview of its health and economic implications. *Economics & Human Biology*, 5 (3): 392–408.

Lingard, J. (2003). A comparative advantage analysis of Kosovan agriculture. *Post-communist Economies*, 15 (3): 417–434.

MAFRD (Ministry of Agriculture, Forestry and Rural Development) (2014). Green Report 2014. Prishtina, Kosovo. Available at http://www.mbpzhr-ks.net/repository/docs/56917_161214_GR_2014_ENG_Final_Printed_Version.pdf.

Malis, C., Rasmussen, E., Poulsen, P., Petersen, I., Christensen, K., Beck-Nielsen, H., Astrup, A. and Vaag, A. (2005). Total and regional fat distribution is strongly influenced by genetic factors in young and elderly twins. *Obesity: A Research Journal*, 13 (12): 2139–2145.

McMichael, P. and Schneider, M. (2011). Food Security Politics and the Millennium Development Goals. *Third World Quarterly*, 32 (1): 119–139.

Moon, W., Florkowski, W. J., Beuchat, L. R., Resurreccion, A. V., Paraskova, P., Jordanov, J., and Chinnan, M. S. (2002). Demand for food variety in an emerging market economy. *Applied Economics*, 34(5): 573–581.

Morales, L.F., Gordon-Larsen, P. and Guilkey, D. (2016). Obesity and health-related decisions: An empirical model of the determinants of weight status across the transition from adolescence to young adulthood. *Economics & Human Biology*, 23: 46–62.

Neufeld, L.M. and Osendarp, S.J.M. (2014). Global, regional and country trends in underweight and stunting as indicators of nutrition and health of populations. In *International Nutrition: Achieving Millennium Goals and Beyond*. Karger Publishers.

Osmani, F., Gorton, M. and White, J. (2013). Agricultural households, poverty and the rural labour market in Kosovo. *Post-Communist Economies*, 25 (2): 241–252.

Pollack, R.A. (1970). Habit formation and dynamic demand functions. *Journal of Political Economy*, 78 (4): 745–763.

Raftopoulou, A. (2016). Geographic determinants of individual obesity risk in Spain: A multilevel Approach. *Economics & Human Biology*, 24: 185–193.

Reinhard, I. and Wijeratne, K.B.S. (2000). *The use of stunting and wasting as indicators for food insecurity and poverty*. Poverty Impact Monitoring Unit, Working Paper 27.

Rosenzweig, M.R. and T.P. Schultz (1982). The behavior of mothers as inputs to child health: The determinants of birth weight, gestation, and rate of fetal growth. In V.R. Fuchs, (Ed.), Economic Aspects of Health, National Bureau of Economic Research, The University of Chicago Press, pp. 53–92.

Stewart, H. and Harris, M. (2005). Obstacles to overcome in promoting dietary variety: The case of vegetables. *Review of Agricultural Economics*, 27 (1): 21–36.

Swinnen, J. and Van Herck, K. (2011). Food security and transition regions. Rome: FAO. Available at http://www.ebrd.com/downloads/research/essay/foodsecurity.pdf.

Theil, H. and Finke, R. (1983). The consumer's demand for diversity. *European Economic Review*, 23 (3): 395–400.

Thiele, S. and Weiss, C. (2003). Consumer demand for food diversity: Evidence for Germany. *Food Policy*, 28 (2): 99–115.

Thomas, D. and Frankenberg, E. (2002). Health, nutrition and prosperity: A microeconomic perspective. *Bulletin of the World Health Organization*, 80 (2): 106–113.

UNDP (United Nations Development Programme) (2014) Kosovo Human Development Report 2014. Available at http://hdr.undp.org/sites/default/files/khdr2014english.pdf.

Victora, C.G. (1992). The association between wasting and stunting: an international perspective. *The Journal of Nutrition*, 122 (5), 1105–1110.

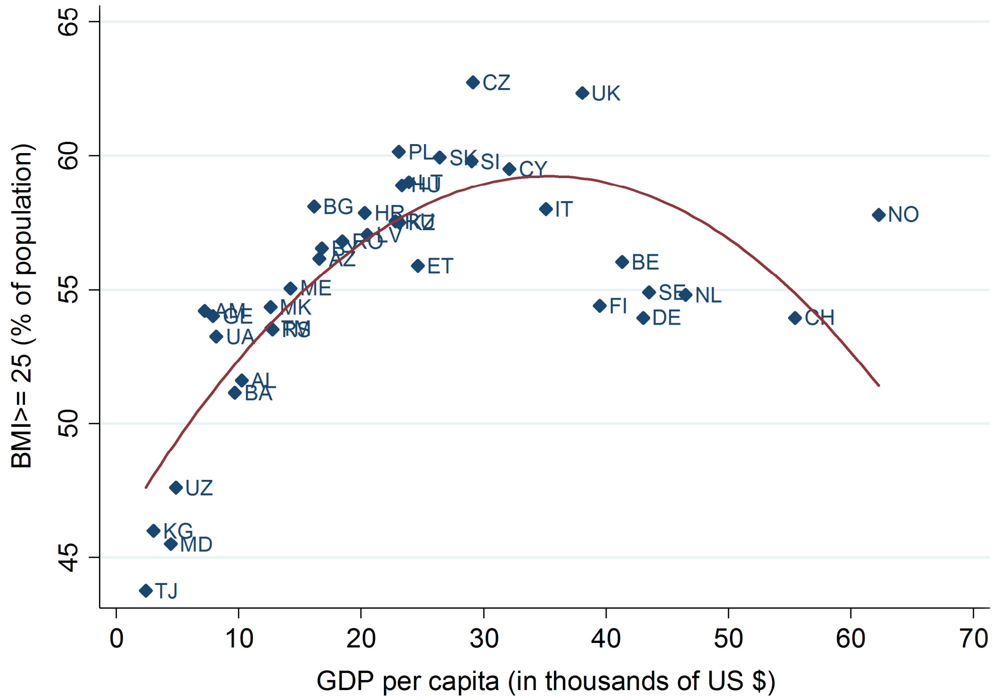
WHO (World Health Organisation) (2014) The International Classification of adult underweight, overweight and obesity according to BMI. Available at http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.

World Bank (2015). Kosovo Country Snapshot. Available at http://www.worldbank.org/content/dam/Worldbank/document/eca/Kosovo-Snapshot.pdf.

World Bank/KAS (2011) Consumption poverty in the Republic of Kosovo in 2009. [Online]. Available at http://siteresources.worldbank.org/INTKOSOVO/Resources/Kosovo_Poverty_for_web_eng.pdf.

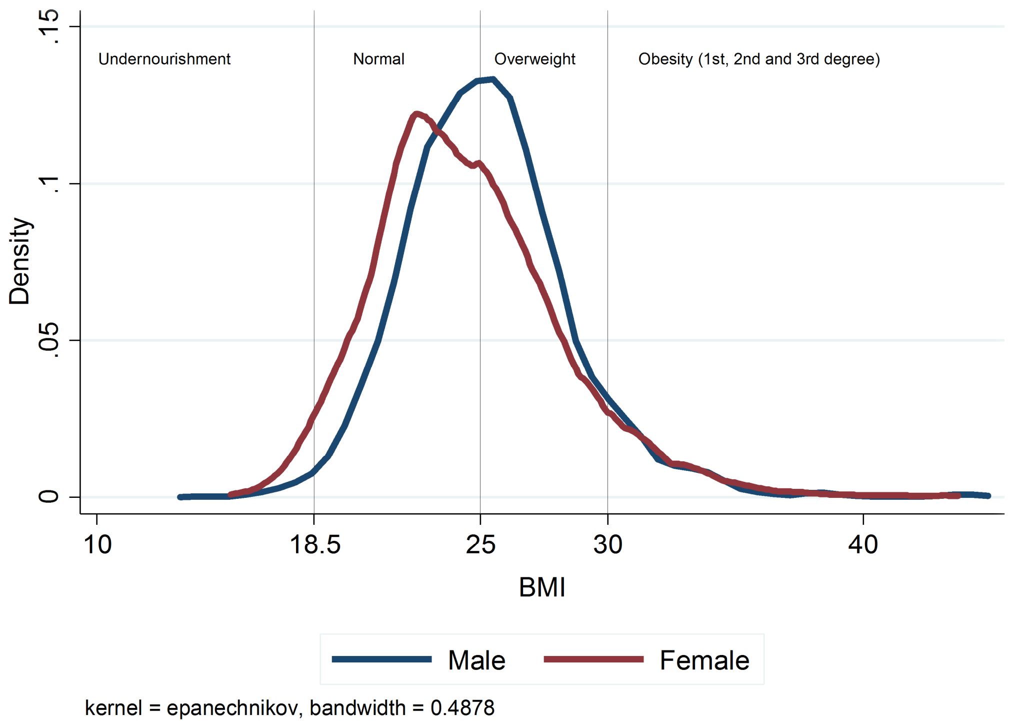
Figures

Figure 1: Scatter plot of overweight and economic development across countries



Source: WHO and World Bank; own processing

Figure 2: Distribution of the BMI of Kosovar adults by gender



Note: Outliers (1%) in the BMI were trimmed.
Source: HBS data of KAS, own processing

Tables

Table 1: Descriptive statistics of variables entering BMI regressions (sample of adults)

Variable	N	Mean	SD	Min	Max
Body mass index	8,902	24.91	3.50	13.73	44.38
Food diversity: Count measure (predicted)	8,739	28.96	6.88	2.15	46.57
Food diversity: Simpson index (predicted)	8,733	0.91	0.03	0.78	0.98
Food diversity: Entropy index (predicted)	8,733	2.87	0.27	1.75	3.55
Dummy: male	8,916	0.50	0.50	0.00	1.00
Dummy: higher education	8,916	0.09	0.29	0.00	1.00
Dummy: secondary education	8,916	0.44	0.50	0.00	1.00
Age	8,916	41.05	17.06	18.00	103.00
Wage (predicted)	8,916	7.31	0.46	5.15	8.63
Propensity to smoke (predicted)	8,916	0.20	0.15	0.00	0.62
Physical activity	8,916	2.22	0.72	1.00	4.00
Dummy: urban area	8,916	0.50	0.50	0.00	1.00
Dummy: area of capital	8,916	0.17	0.37	0.00	1.00

Note: Summary statistics for adults sample.

Table 2: QR estimates of BMI's determinants (predicted Count Measure)

Variable	Q=0.10	Q=0.50	Q=0.90
Food diversity: Count measure (predicted)	0.002***	-0.000	-0.001*
	(0.000)	(0.000)	(0.001)
Dummy: male	0.083***	0.072***	0.080***
	(0.010)	(0.006)	(0.013)
Dummy: higher education	0.036*	-0.039***	0.002
	(0.021)	(0.013)	(0.026)
Dummy: secondary education	0.015	-0.015**	-0.000
	(0.010)	(0.006)	(0.013)
Age	0.018***	0.015***	0.021***
	(0.002)	(0.001)	(0.002)
Age squared	-0.017***	-0.012***	-0.019***
	(0.002)	(0.001)	(0.002)
Wage (predicted)	-0.045**	0.014	-0.034
	(0.020)	(0.013)	(0.026)
Dummy: propensity to smoke (predicted)	-0.138***	-0.158***	-0.197***
	(0.034)	(0.021)	(0.043)
Physical activity	0.007**	0.011***	0.004
	(0.003)	(0.002)	(0.004)
Dummy: urban area	-0.018***	0.001	0.018***
	(0.005)	(0.003)	(0.007)
Dummy: area of capital	0.021***	-0.001	-0.019**
	(0.006)	(0.004)	(0.008)
Constant	2.923***	2.770***	3.135***
	(0.105)	(0.066)	(0.132)
Pseudo R^2	0.140	0.153	0.125
N	8,725	8,725	8,725

Note: The dependent variable is log(BMI). Robust standard errors are presented in parentheses. Quantile regressions estimated on the sample of adults. Primary education is the reference category for the education dummy set. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: QR estimates of BMI's determinants (predicted Simpson index)

Variable	Q=0.10	Q=0.50	Q=0.90
Food diversity: Simpson index (predicted)	0.373***	-0.062	-0.263**
	(0.100)	(0.061)	(0.122)
Dummy: male	0.077***	0.071***	0.079***
	(0.010)	(0.006)	(0.013)
Dummy: higher education	0.017	-0.041***	0.010
	(0.021)	(0.013)	(0.025)
Dummy: secondary education	0.005	-0.015**	0.003
	(0.010)	(0.006)	(0.012)
Age	0.017***	0.015***	0.022***
	(0.002)	(0.001)	(0.002)
Age squared	-0.015***	-0.012***	-0.019***
	(0.002)	(0.001)	(0.002)
Wage (predicted)	-0.021	0.015	-0.040*
	(0.020)	(0.012)	(0.024)
Dummy: propensity to smoke (predicted)	-0.132***	-0.157***	-0.188***
	(0.035)	(0.021)	(0.042)
Physical activity	0.005	0.010***	0.002
	(0.003)	(0.002)	(0.004)
Dummy: urban area	-0.016***	0.002	0.018***
	(0.005)	(0.003)	(0.006)
Dummy: area of capital	0.015**	-0.000	-0.016**
	(0.006)	(0.004)	(0.008)
Constant	2.513***	2.814***	3.385***
	(0.109)	(0.066)	(0.132)
Pseudo R^2	0.139	0.153	0.125
N	8,719	8,719	8,719

Note: The dependent variable is log(BMI). Robust standard errors are presented in parentheses. Quantile regressions estimated on the sample of adults. Primary education is the reference category for the education dummy set. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: QR estimates of BMI's determinants for adults (predicted Entropy index)

Variable	Q=0.10	Q=0.50	Q=0.90
Food diversity: Entropy index (predicted)	0.048***	-0.007	-0.029**
	(0.011)	(0.007)	(0.014)
Dummy: male	0.081***	0.071***	0.081***
	(0.010)	(0.006)	(0.013)
Dummy: higher education	0.024	-0.041***	0.007
	(0.021)	(0.013)	(0.026)
Dummy: secondary education	0.009	-0.015**	0.002
	(0.010)	(0.006)	(0.012)
Age	0.017***	0.015***	0.022***
	(0.002)	(0.001)	(0.002)
Age squared	-0.016***	-0.012***	-0.019***
	(0.002)	(0.001)	(0.002)
Wage (predicted)	-0.031	0.016	-0.038
	(0.020)	(0.012)	(0.025)
Dummy: propensity to smoke (predicted)	-0.142***	-0.156***	-0.196***
	(0.035)	(0.021)	(0.043)
Physical activity	0.006*	0.010***	0.003
	(0.003)	(0.002)	(0.004)
Dummy: urban area	-0.017***	0.002	0.020***
	(0.005)	(0.003)	(0.007)
Dummy: area of capital	0.018***	-0.001	-0.018**
	(0.006)	(0.004)	(0.008)
Constant	2.769***	2.774***	3.216***
	(0.100)	(0.061)	(0.123)
Pseudo R^2	0.140	0.153	0.125
N	8,719	8,719	8,719

Note: The dependent variable is log(BMI). Robust standard errors are presented in parentheses. Quantile regressions estimated on the sample of adults. Primary education is the reference category for the education dummy set. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendices

Table A.1: Descriptive statistics of variables entering diet diversity regressions

Variable	N	Mean	SD	Min	Max
Food diversity: Count measure	8,916	28.76	11.31	0.00	66.00
Food diversity: Simpson index	8,910	0.91	0.07	0.00	0.98
Food diversity: Entropy index	8,910	2.86	0.48	0.00	3.94
Household income	8,916	5521.52	4335.16	0.00	57700.00
Own agricultural consumption to household income ratio	8,739	0.02	0.05	0.00	0.94
FAFH to household income ratio	8,739	0.01	0.04	0.00	1.00
Dummy: 1st quarter	8,916	0.25	0.43	0.00	1.00
Dummy: 2nd quarter	8,916	0.26	0.44	0.00	1.00
Dummy: 3rd quarter	8,916	0.25	0.43	0.00	1.00
Dummy: 4th quarter	8,916	0.25	0.43	0.00	1.00
Dummy: region of Gjakova	8,916	0.14	0.34	0.00	1.00
Dummy: region of Gjilani	8,916	0.13	0.34	0.00	1.00
Dummy: region of Mitrovica	8,916	0.13	0.34	0.00	1.00
Dummy: region of Peja	8,916	0.14	0.35	0.00	1.00
Dummy: region of Prizren	8,916	0.16	0.36	0.00	1.00
Dummy: region of Prishtina	8,916	0.17	0.37	0.00	1.00
Dummy: region of Ferizaj	8,916	0.14	0.34	0.00	1.00
Dummy: male	8,916	0.50	0.50	0.00	1.00
Dummy: higher education	8,916	0.09	0.29	0.00	1.00
Dummy: secondary education	8,916	0.44	0.50	0.00	1.00
Age	8,916	41.05	17.06	18.00	103.00
Dummy: employed	8,916	0.89	0.31	0.00	1.00
Household size	8,916	6.67	3.44	1.00	29.00
Dummy: having children (aged below 7)	8,916	0.42	0.49	0.00	1.00
Dummy: urban area	8,916	0.50	0.50	0.00	1.00

Note: Summary statistics for the adults sample.

Table A.2: Descriptive statistics of variables entering Heckman selection model for wages

Variable	N	Mean	SD	Min	Max
Wage	2,018	7.97	0.65	3.93	10.59
Dummy: male	8,916	0.50	0.50	0.00	1.00
Dummy: higher education	8,916	0.09	0.29	0.00	1.00
Dummy: secondary education	8,916	0.44	0.50	0.00	1.00
Age	8,916	41.05	17.06	18.00	103.00
Dummy: 1st quarter	8,916	0.25	0.43	0.00	1.00
Dummy: 2nd quarter	8,916	0.26	0.44	0.00	1.00
Dummy: 3rd quarter	8,916	0.25	0.43	0.00	1.00
Dummy: 4th quarter	8,916	0.25	0.43	0.00	1.00
Dummy: region of Gjakova	8,916	0.14	0.34	0.00	1.00
Dummy: region of Gjilani	8,916	0.13	0.34	0.00	1.00
Dummy: region of Mitrovica	8,916	0.13	0.34	0.00	1.00
Dummy: region of Peja	8,916	0.14	0.35	0.00	1.00
Dummy: region of Prizren	8,916	0.16	0.36	0.00	1.00
Dummy: region of Prishtina	8,916	0.17	0.37	0.00	1.00
Dummy: region of Ferizaj	8,916	0.14	0.34	0.00	1.00
Dummy: urban area	8,916	0.50	0.50	0.00	1.00
Dummy: rent, dividends, interest	8,916	0.01	0.09	0.00	1.00
Dummy: social welfare benefits	8,916	0.02	0.14	0.00	1.00
Dummy: pensions from Kosovo	8,916	0.11	0.32	0.00	1.00
Dummy: pensions from outside of Kosovo	8,916	0.01	0.12	0.00	1.00
Dummy: cash remittances from Kosovo	8,916	0.00	0.06	0.00	1.00
Dummy: cash sent from abroad by current HH members	8,916	0.01	0.08	0.00	1.00
Dummy: cash sent from abroad by relatives and other persons	8,916	0.04	0.20	0.00	1.00
Dummy: gifts in kind from abroad	8,916	0.00	0.04	0.00	1.00
Dummy: male	8,916	0.50	0.50	0.00	1.00
Dummy: higher education	8,916	0.09	0.29	0.00	1.00
Dummy: secondary education	8,916	0.44	0.50	0.00	1.00
Household size	8,916	6.67	3.44	1.00	29.00
Dummy: having children (aged below 7)	8,916	0.42	0.49	0.00	1.00

Note: Summary statistics for the adults sample.

Table A.3: Descriptive statistics of variables entering Probit model for propensity to smoke

Variable	N	Mean	SD	Min	Max
Dummy: smoking	8,916	0.20	0.40	0.00	1.00
Individual income	8,916	1293.54	2240.57	0.00	57700.00
Dummy: male	8,916	0.50	0.50	0.00	1.00
Dummy: employed	8,916	0.89	0.31	0.00	1.00
Dummy: higher education	8,916	0.09	0.29	0.00	1.00
Dummy: secondary education	8,916	0.44	0.50	0.00	1.00
Age	8,916	41.05	17.06	18.00	103.00
Physical activity	8,916	2.22	0.72	1.00	4.00
Dummy: 1st quarter	8,916	0.25	0.43	0.00	1.00
Dummy: 2nd quarter	8,916	0.26	0.44	0.00	1.00
Dummy: 3rd quarter	8,916	0.25	0.43	0.00	1.00
Dummy: 4th quarter	8,916	0.25	0.43	0.00	1.00
Dummy: region of Gjakova	8,916	0.14	0.34	0.00	1.00
Dummy: region of Gjilani	8,916	0.13	0.34	0.00	1.00
Dummy: region of Mitrovica	8,916	0.13	0.34	0.00	1.00
Dummy: region of Peja	8,916	0.14	0.35	0.00	1.00
Dummy: region of Prizren	8,916	0.16	0.36	0.00	1.00
Dummy: region of Prishtina	8,916	0.17	0.37	0.00	1.00
Dummy: region of Ferizaj	8,916	0.14	0.34	0.00	1.00
Dummy: urban area	8,916	0.50	0.50	0.00	1.00

Note: Summary statistics for the adults sample.

Table A.4: OLS estimates of food diversity

Variable	(CM)	(SI)	(EI)
Household income (log)	4.442***	0.015***	0.161***
	(0.329)	(0.002)	(0.015)
Own agricultural consumption to household income ratio	-6.841	-0.039	-0.321
	(4.199)	(0.061)	(0.259)
FAFH to household income ratio	-5.635	-0.030	-0.223
	(3.988)	(0.039)	(0.241)
Dummy: 1st quarter	-0.886	-0.008*	-0.061**
	(0.587)	(0.004)	(0.026)
Dummy: 2nd quarter	0.062	0.001	0.004
	(0.592)	(0.004)	(0.028)
Dummy: 3rd quarter	-0.114	0.004	0.008
	(0.601)	(0.004)	(0.026)
Dummy: male	-0.406***	-0.001	-0.014**
	(0.145)	(0.001)	(0.007)
Dummy: higher education	2.047***	0.006**	0.073***
	(0.520)	(0.003)	(0.021)
Dummy: secondary education	0.953***	0.003	0.038***
	(0.298)	(0.002)	(0.014)
Age	0.044	0.000	0.002
	(0.038)	(0.000)	(0.002)
Age squared x 10^{-2}	-0.054	-0.000	-0.003
	(0.047)	(0.000)	(0.002)
Dummy: employed	0.138	-0.006	-0.014
	(0.642)	(0.004)	(0.028)
Household size	-0.048	-0.002**	-0.009
	(0.117)	(0.001)	(0.006)
Dummy: having children (aged below 7)	0.815	0.005	0.036
	(0.500)	(0.003)	(0.023)
Dummy: urban area	6.489***	0.023***	0.240***
	(0.434)	(0.003)	(0.019)
Constant	-14.774***	0.769***	1.313***
	(2.793)	(0.021)	(0.129)
Regional control variables	YES	YES	YES
R^2	0.374	0.187	0.317
N	8,739	8,733	8,733

Note: Standard errors clustered at household level are presented in parentheses. OLS diet diversity regressions estimated on the sample of adults. Dummy variables for primary education, 4th quarter, and region of Ferizaj are the reference categories for the respective dummy sets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Probit estimates of propensity to smoke

Variable	(1)
Individual income (log)	0.046***
	(0.006)
Dummy: male	0.799***
	(0.041)
Dummy: employed	0.125
	(0.102)
Dummy: higher education	-0.346***
	(0.065)
Dummy: secondary education	-0.006
	(0.040)
Age	0.095***
	(0.008)
Age squared	-0.096***
	(0.009)
Physical activity=2	-0.086
	(0.063)
Physical activity=3	-0.145*
	(0.076)
Physical activity=4	-0.080
	(0.082)
Dummy: 1st quarter	-0.001
	(0.048)
Dummy: 2nd quarter	0.087*
	(0.047)
Dummy: 3rd quarter	-0.019
	(0.047)
Dummy: urban area	0.122***
	(0.034)
Constant	-3.489***
	(0.152)
Regional control variables	YES
Pseudo R^2	0.149
N	8,916

Note: Robust standard errors are presented in parentheses. Probit model of smoking estimated on the sample of adults. Dummy variables for primary education, low physical activity (=1), 4th quarter, and region of Ferizaj are the reference categories for the respective dummy sets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6 Estimates of Heckman selection model for wages

Variable	Selection equation	Regression equation
Dummy: Rent, dividends, interest	-0.853***	
	(0.193)	
Dummy: Social welfare benefits	-0.720***	
	(0.131)	
Dummy: Pensions from Kosovo	-1.843***	
	(0.234)	
Dummy: Pensions from outside of Kosovo	-1.289***	
	(0.350)	
Dummy: Cash remittances from Kosovo	-1.040**	
	(0.415)	
Dummy: Cash sent from abroad by current HH members	-0.695***	
	(0.238)	
Dummy: Cash sent from abroad by relatives and other persons	-0.924***	
	(0.109)	
Dummy: Gifts in kind from abroad	-0.746	
	(0.627)	
Household size	-0.035***	
	(0.006)	
Dummy: having children (aged below 7)	-0.168***	
	(0.040)	
Dummy: male	1.118***	0.180**
	(0.038)	(0.072)
Dummy: higher education	1.112***	0.992***
	(0.059)	(0.075)
Dummy: secondary education	0.378***	0.448***
	(0.041)	(0.042)
Age	0.142***	0.073***
	(0.009)	(0.012)
Age squared	-0.158***	-0.077***
	(0.011)	(0.014)
Dummy: 1st quarter	-0.165***	-0.177***
	(0.050)	(0.039)
Dummy: 2nd quarter	0.018	-0.076**
	(0.049)	(0.037)
Dummy: 3rd quarter	-0.038	-0.075**
	(0.049)	(0.037)
Dummy: urban area	0.200***	0.061**
	(0.036)	(0.030)
Constant	-4.239***	5.409***
	(0.184)	(0.381)
Regional dummy variables	YES	YES
Inverse mills ratio (lambda)	0.306***	
	(0.087)	
Rho	0.500	
Sigma	0.625	
Number of individuals	8,916	
Number, censored	6,898	

Note: Robust standard errors are presented in parentheses. Heckman selection model estimated on the sample of adults. Dummy variables for primary education, 4th quarter, and region of Ferizaj are the reference categories for the respective dummy sets. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.