NBER WORKING PAPER SERIES

ESTIMATION OF A DYNAMIC MODEL OF WEIGHT

Shu Wen Ng
Edward C. Norton
David K. Guilkey
Barry M. Popkin

Working Paper 15864
http://www.nber.org/papers/w15864

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2010

This research uses data from China Health and Nutrition Survey (CHNS). We thank the Institute of Nutrition and Food Safety, China Center for Disease Control and Prevention, Carolina Population Center, at the University of North Carolina at Chapel Hill, the NIH (R01-HD30880, DK056350, and R01-HD38700) and the Fogarty International Center for financial support for the CHNS data collection and analysis files, and for partial support of Ng and Popkin. We also thank Harsha Thirumurthy, Shoou-Yih Daniel Lee, Linda Adair, William Greene, Trevor Long, participants of the Annual Health Econometrics Workshop, seminar participants at the University of Michigan, and three anonymous reviewers for helpful comments and suggestions for improvements. None of the authors has conflict of interests of any type with respect to this manuscript. All errors are of the authors' alone. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Estimation of a Dynamic Model of Weight
Shu Wen Ng, Edward C. Norton, David K. Guilkey, and Barry M. Popkin
NBER Working Paper No. 15864
April 2010
JEL No. I12

ABSTRACT

The ongoing debate about the economic causes of obesity has focused on the changing relative prices of diet and exercise. This paper uses a model that explicitly includes time and spatially varying community-level urbanicity and price measures as instruments to obtain statistically correct measures for the endogenous effects of diet, physical activity, drinking, and smoking on weight. We apply a dynamic panel system GMM estimation model to longitudinal (1991–2006) data from China to model weight and find that among adult men in China, about 6.1% of weight gain was due to declines in physical activity and 2.9-3.8% was due to dietary changes over this period. In the long run, physical activity can account for around 6.9% of weight gain, while diet can account for 3.2-4.2% of weight gain.

Shu Wen Ng
Department of Nutrition
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599
shuwen@email.unc.edu

Edward C. Norton
Department of Health Management and Policy
Department of Economics
University of Michigan
School of Public Health
109 S. Observatory Drive, M3108 SPH II
Ann Arbor, MI 48109-2029
and NBER
ecnorton@umich.edu

David K. Guilkey
Department of Economics
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599
david_guilkey@unc.edu

Barry M. Popkin
Department of Economics
Department of Nutrition
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599
popkin@unc.edu

I. Introduction

There is an ongoing debate among economists about the economic causes of obesity, in particular the relative importance of diet and exercise. Some empirical work focuses on the rise in the relative price of physical activity (Pratt, Macera et al., 2004; Sturm, 2005) both because physical activity has become less a part of daily activities and because the value of time has risen with income. Others point to the drop in the relative price of calories (Chou, Grossman et al., 2004; Cutler, Glaeser et al., 2003; Drewnowski & Darmon, 2005; Kuchler, Tegene et al., 2004; Schroeter, Lusk et al., 2008) due to higher supply from agricultural innovations and improvements in the preserving, packaging, preparation, and transportation of food. Others model the change in both relative prices of physical activity and caloric intake (Cawley, 2004; Lakdawalla & Philipson, 2009; Lakdawalla, Philipson et al., 2005; Philipson & Posner, 2003; Rashad, 2006; Rashad & Grossman, 2004). However, none of these economic studies have come up with a definitive way to quantify the contribution of diet versus physical activity while controlling for weight-related health behaviors like smoking and drinking. Part of the difficulty stems from the lack of longitudinal individual-level data and the difficulty in using cross sectional data in an empirical examination of the forces contributing to weight over time. Clinical studies, on the other hand, tend to focus specifically only on one factor at a time (Brien, Katzmarzyk et al., 2007; Taylor, Jatulis et al., 1994), and so cannot resolve the debate.

Understanding which affects weight more — diet or physical activity — has great implications for both public policy and for individuals. For policy, it would help us understand which public policies such as taxing fat, subsidizing vegetables, and promoting green spaces would be most effective. For those who want to manage or lower their weight, it would help them understand where to focus their efforts.

At first glance, modeling an individual's weight appears straightforward: define the number of calories consumed and expended, and determine the resulting trends in weight. However, the complex relationships among physical activity, diet, drinking, smoking, and weight are lost in a simplistic formulation (Moore, 2000; Prentice & Jebb, 2004) due to the fact that all of these variables are choice variables to the individual which results in endogeneity problems. For example, lower physical activity and increased sedentary behaviors among those who are heavy may be the consequence of being heavy (e.g., social stigma, being shunned from sports, or being physically unable to partake in activity). Similarly, it is possible that people consume more (as a coping mechanism) in reaction to being marginalized due to their weight. In addition, endogeneity bias may arise due to omitted variables (such as genetic endowments) are determinants of both an included explanatory variables and weight.

Many previous studies typically only look at the relationships between physical activity and weight, and diet and weight separately, without considering the endogenous decisions of contemporaneous and lagged diet, physical activities, smoking and drinking on weight. Moreover, calories from fat versus calories from carbohydrates or proteins, or calories from animal source foods versus fruits and vegetables might affect weight differently (Miller, Lindeman et al., 1990; Tryon, 1987), and diet, physical activity, smoking and drinking combined can interact to affect weight (Astrup, 1999; Klesges, Meyers et al., 1989). For example, people with low physical activity levels but high fat intake have slower metabolisms, which results in greater weight gain (Bray & Popkin, 1998), particularly for people in developing countries who might have experienced undernutrition during prenatal and postnatal growth (Frisancho, 2003; James & Ralph, 1999). Smoking has also been found to affect both the metabolic mechanism and food preferences, and thus affect weight (Klesges, Meyers et al., 1989). If the endogeneity is not corrected, the results will be inconsistent.

To account for potential endogeneity, economists often employ instrumental variable (IV) techniques, longitudinal fixed effects (FE) models or combine both these approaches (IV-FE). These methods have the potential to generate consistent estimates if reliable instruments are available in the data. There is also the issue of serially correlated errors due to time invariant unobserved heterogeneity, which can result in incorrect standard errors.

Beyond the issues of endogeneity and serially correlated errors however, there is the added complexity of autocorrelation of diet, physical activities, smoking and drinking decisions, and because past weight will be associated with current weight. If the correlation of weight over time is not controlled for (via the inclusion of lagged weight), then the estimated effect of past weight will tend to be too large as well as inefficient (large standard errors).

We estimate the relative importance of the diet, physical activity, and the health behaviors of smoking and drinking on weight among adult men over a period of rapid economic growth in China by employing two strategies. First, we use a model that explicitly includes time and spatially varying macro-level factors such as urbanicity and prices to be used as instruments to correct for the endogenous micro-level choices of diet, physical activity and other health behaviors that affect weight over time. Second, we apply a dynamic panel system generalized methods of moments (GMM) model estimation model, which allows current weight to depend on prior weight and endogenous decisions about physical diet, drinking, smoking and physical activity. This estimation approach uses statistical methods that control for the endogeneity problems, the related correlation of errors for the same individual through time, and explicitly models lagged weight within the model.

We estimate our dynamic panel model using a GMM estimator developed by Blundell and Bond (1998) that exploits a large set of moment conditions and combines in a system, the regression-in-differences with the regression-in-levels models. We provide a comparison of these results to those derived from models that assume exogeneity to show how the failure to correct for endogeneity and temporally correlated errors can affect the findings. The coefficients from the dynamic model will show the relative strength of diet and physical activity on weight, and these results can be combined with known changes in types of diet and levels of physical activity to show which factors affected weight gain the most, at least among adult men in China. We use six waves of the longitudinal China Health and Nutrition Surveys (CHNS) that contain detailed individual-level data on anthropometrics, dietary consumption, energy expenditure, as well as time varying community measures of urbanicity and prices that can be used as instrumental variables for potentially endogenous variables. We found that declines in physical activities, increases in fat, decreases in carbohydrates, and increases in oils and fats as a proportion of a person's dietary intake are positively associated with weight, and the coefficients of these endogenous variables from the GMM model are larger and more significant compared to models without corrections for endogeneity. Calculations show that on average, 6.1% of weight gain among adult men in China from 1991 to 2006 was due to declines in physical activity, while 2.9 to 3.8% was due to dietary changes. In the long run, physical activity can account for around 6.9% of weight gain, while diet can account for around 3.2 to 4.2% of weight gain.

Globally, the growing epidemic of overweight and obesity, risk markers for a large number of chronic diseases, will have severe consequences on its economic productivity and will become a significant health care burden. For China, it is estimated that the total direct (health care) and indirect (disability, mortality, and morbidity) cost of overweight and related diseases was 3.5% of China's gross national product (GNP) in 2000, but will grow to almost 9% by 2025 (Popkin, Kim et al., 2006).

II. Modeling the dynamics and determinants of weight

A theoretical model provides guidance on which variables are endogenous and which other variables are potential instruments. Our theoretical model is based on work by others (Cawley, 2004; Chou, Grossman et al., 2004; Drewnowski & Darmon, 2005; Lakdawalla, Philipson et al., 2005; Philipson & Posner, 2003; Rashad & Grossman, 2004), and stems from a rational choice model. As these economists have noted, this model is not meant to be an introspective guide to how people think about their choices, but rather an internally-consistent methodology to generate hypotheses about their behavior.

A. *Dynamics of weight*

An individual's utility in current period, t, depends on food consumption, F, physical activity (A), other health behaviors (such as smoking), other consumption, C, and current weight, W. Utility U increases with consumption of food, physical activity, other health behaviors and other consumption, but is increasing in weight only if current weight is less than ideal weight, $\ddot{W}$. Otherwise utility declines with weight. The marginal utility of eating decreases as weight exceeds ideal weight, because eating increases weight. The assumption is that there is an ideal weight, $\ddot{W}$, holding other consumption constant. In other words, $\ddot{W}$ is the weight that would be chosen if achieving one's preferred weight were costless. This subjective ideal weight may or may not correspond to the weight that maximizes health or longevity, although it is likely to be influenced by concern with these factors. But the ideal weight in this model is not necessarily the preferred weight in the economic sense because it does not consider the full range of costs and benefits of achieving it. In other words, a person's rationally chosen weight is the one that makes him the happiest given the existing costs and benefits of food consumption, physical activity and other consumption.

Because this model focuses on weight, we conceptualize food consumption simply as caloric intake, including calories from alcohol. Two other behaviors affect weight — physical activity, A, and smoking, S. Both affect utility directly and indirectly (as determinants of weight).

An individual's physical activity level depends on the level of development, D, where she lives, such that $A_t = A(D_t)$.

(1) $$W_t = (1-\delta)W_{t-1} + g(F_{t-1}, S_{t-1}, A(D_{t-1})),$$

where $\delta < 1$ and g is continuous, concave, increasing in food or alcohol consumption, decreasing in physical activity level, and decreasing in smoking level ($g_F \geq 0$, $g_A \leq 0$ and $g_S \leq 0$).

Individuals are subject to a budget constraint each period: $p_F F + p_S S + p_c C \leq I$, where p_F, p_S and p_c are the prices of food (including alcohol), cigarettes, and other consumption goods respectively, and I is income. Consistent with existing literature, this does not account for borrowings and savings over time.

When maximizing utility, an individual chooses F_{t-1}, S_{t-1}, and A_{t-1} simultaneously. These choices are endogenous to each other because of implicit tradeoffs in terms of time and money spent on each. In addition, the choice variables are serially correlated because of habit formation, addiction (especially smoking and alcohol), unobserved preferences, genetics, and shared environmental factors. This suggests that it is important to control for the endogeneity bias from diet, smoking, and physical activities choices on weight, and for the serial correlation of these decisions over time.

B. *Steady State Determinants of Weight, Diet, Smoking and Physical Activity*

This dynamic maximization problem yields a unique steady-state in weight, food consumption, smoking, physical activity and weight (see the Appendix) determined by income, I, food (and alcohol) prices, p_F, cigarette price, p_S, and urbanicity, D, such that $W^*(I, p_F, p_S, D)$, $F^*(I, p_F, p_S, D)$, $S^*(I, p_F, p_S, D)$, $A^*(I, p_F, p_S, D)$. If these factors are exogenous to weight, diet, smoking, and physical activity, and vary over space and time, then they would make ideal instruments to correct for bias caused by the inclusion of endogenous explanatory variable.

Increases in income raise weight at low levels of income, but at high levels of income, further increases could lower weight (i.e., W_I^* has an inverted U-shape). An increase in income lowers the marginal cost of spending on weight gain (food and alcohol consumption), but also affects the marginal value of weight. Income is also related to one's physical activity level, because A is a function of job characteristics. In a country like China, those who are poor generally have jobs that require greater physical activity, so we assume that $A_I < 0$. For people who are underweight, a rise in income will typically increase weight both through greater food consumption and less physical activity on the job. For people who are overweight, an increase in income may eventually lead to enough resources to reduce their weight.

Increasing the price of food and alcohol, p_F, raises the marginal cost of caloric intake, so food and alcohol consumption decreases, so that $F_{pF}^*(I, p_F, p_S, D) < 0$. The decrease in food and alcohol consumption will also lower weight, so that $W_{pF}^*(I, p_F, p_S, D) < 0$. Increasing the price of cigarettes, p_S, decreases smoking (i.e., $S_{pS}^*(I, p_F, p_S, D) < 0$), which may alter metabolic mechanisms that results in weight gain (Klesges, Meyers et al., 1989), so that $W_{pS}^*(I, p_F, p_S, D) > 0$. Hence, prices are important determinants of weight and are exogenous factors that need to be included in any model of weight.

Community-level urbanicity, D is exogenous to individual choice assuming that people who move do not do so based primarily on these community-level characteristics. Development affects prices of food, cigarettes, other consumption goods, and income, such that increased development lower prices and raise incomes. Hence it can be thought of as an argument for p_F, p_S and I. Using chain rule, the effect of urbanicity on food consumption, smoking and physical activity levels are $F_D^*>0$, $S_D^*>0$ and $A_D^*<0$. Urbanicity lowers physical activity at work, access to technologies that aid work and domestic activities, and the availability of motorized transportation. Also, one would expect urbanicity at the community-level to reduce food and cigarette prices through lowering transportation costs, and lessening the time involved in purchasing these items.

We expect past caloric intake, dietary fat intake, and alcohol consumption to be positively related to current weight, and past levels of physical activity and smoking levels to be negatively related to current weight. Moreover, past clinical studies suggest that physical activity may be more important than diet in weight control (King, Frey-Hewitt et al., 1989) due to relative ease of communicating its benefits, and the resultant metabolic effects on lipid mobilization, oxidation and biochemical changes, which help improved regulation of body weight (Saris, 1998). Therefore, if past diet measures and physical activity factors are statistically significant in explaining weight, we can determine the relative contribution of each of these, and hence inform on public policy and individual choices.

III. Empirical Modeling

The dynamic empirical model that relates weight to its own lagged value along with lagged food (and alcohol) consumption, lagged physical activity, and lagged smoking status takes the following form:

(2) $$W_{it} = \alpha W_{i,t-1} + \beta F_{i,t-1} + \theta S_{i,t-1} + \gamma A_{i,t-1} + \pi X_{it} + \eta_i + \mu_{it},$$

where W_{it} denotes weight in the current wave t for individual i; $W_{i,t-1}$ denotes weight in the prior wave for individual i; $F_{i,t-1}$ denote lagged values of two sets of consumption variables, the first set includes total caloric intake, energy from dietary fat, energy from carbohydrates and drinking status, while the alternative set includes total caloric intake, energy from animal source foods, energy from fruits and vegetables, energy from edible oils and fats, energy from whole grain sources, energy from rice, and drinking status ; $S_{i,t-1}$ denote lagged smoking status; $A_{i,t-1}$ denote lagged total physical activities; X_{it} denotes other control variables such as age, marital status, educational attainment, predicted household income and time dummies; α, β, θ, γ and π denote the vectors of coefficients for the explanatory variables; η_i denotes unobserved time invariant individual characteristics, and μ_{it} denotes a time varying disturbance term.

We expect β_{kcal}, β_{efat} , β_{ecarb} and β_{drink} (as well as β_{kcal}, $\beta_{eanimal}$, $\beta_{oilfats}$ and β_{drink}) to be positively related to W_{it} ; θ (the coefficient for lagged current smoking status) to be negatively related to W_{it}; and γ (the coefficient for lagged physical activity) to be negatively related to W_{it}. If we find that β_{kcal} , β_{efat}, β_{ecarb}, β_{drink} , θ and γ to be statistically significant, then it we can determine the contribution of caloric intake, dietary fat intake, carbohydrates, drinking, smoking and physical activity in determining weight and from that tell which of these are the most important in affecting weight gain.

To determine the type of estimation method to use, it is important to discuss the assumptions made about:

1) The correlation between explanatory variables and η_i
2) Autocorrelation: correlation in the time varying error terms over time (e.g., corr($\mu_{i,t-1}$, μ_{it}))
3) The type of correlation between the explanatory variables and μ_{it} , μ_{it-1} or μ_{it+1}

It is clear from the dynamic form of the stochastic model that at a minimum lagged weight will be correlated with η_i, the time invariant error term and it is also highly likely that there will be overlap in the set of unobserved fixed characteristics of the individuals that affect weight, diet, physical activity, smoking and drinking that will cause correlation between η_i and these variables as well. First differencing will drop η_i along with all time invariant observed variables from the model:

(3) $$\Delta W_{it} = \alpha \Delta W_{i,t-1} + \beta \Delta F_{i,t-1} + \theta \Delta S_{i,t-1} + \gamma \Delta A_{i,t-1} + \pi \Delta X_{it} + \Delta \mu_{it} .$$

We assume that the time-varying error is not correlated with the explanatory variables, which means that in differenced form food and alcohol consumption, smoking, and physical activity, are uncorrelated with the error term in equation (3). Differenced weight may still be correlated with the differenced error term, implying that OLS estimation will be inconsistent. However, $W_{i,t-2}$ will not be, and can therefore be used as an instrument. Still, this instrumental variables estimation in differences tends to yield imprecise parameter estimates if α is large (Alonso-Borrego & Arellano, 1999; Blundell & Bond, 1998).

An alternative (Blundell & Bond, 1998) is to estimate the model in levels, with $\Delta W_{i,t-2}$ used as an instrument for $W_{i,t-1}$ in equation (2). This method, of course, must assume that there is no correlation between the other explanatory variables and either the time invariant or time varying error term. A more efficient estimator, (Blundell & Bond, 1998) would jointly estimate equations (2) and (3) using a system GMM approach. The system GMM estimator uses lagged first difference as instrument for equations in levels as well as the lag variable instruments for first difference equations. With a panel, we can derive a set of instruments which are both correlated with $\Delta W_{i,t-1}$ and orthogonal to $\Delta \mu_{i,t-1}$. For instance, in the absence of serial autocorrelation, the lagged level $W_{i,t-2}$ will be correlated with $\Delta W_{i,t-1}$ but uncorrelated with $\Delta \mu_{i,t-1}$. Each additional time period can add additional valid instruments. This can be similarly done for any endogenous explanatory variables in the model, giving rise to an instrument matrix denoted $Z_i =(W_{,it,}, F_{it}, S_{it}, A_{it}$ and $X_{it})$. The moment conditions are thus:

(4) $$E[\Delta Z_{it}^{'} \Delta \mu_{it}] = 0 \text{, where } \Delta \mu_{it} = (\Delta \mu_{i3}, \Delta \mu_{i4}, \ldots, \Delta \mu_{iT})'$$

As already noted, it is highly likely that there will be correlation between diet, physical activity, smoking, and drinking and the time invariant error and so instruments are needed for these variables in addition to lagged weight in equation (2). It is also possible that there will be correlation between these variables and the time varying error term, meaning that instruments may be needed for these variables even in differenced form in equation (3). Autocorrelation in the time varying error could also invalidate $W_{i,t-2}$ as an instrument in equation (3). The validity of the estimation here therefore, rests on the assumption that the ($W_{,it,}$, F_{it} , S_{it} , A_{it} and X_{it}) series each satisfy a mean stationarity assumption, yielding the additional moment conditions:

(5) $$E[\Delta Z_{i,t-1}(\eta_i + \mu_{it})] = 0$$

Separate instrument sets must be specified for equations (2) and (3) in the system GMM approach, and we discuss these sets further below. Cameron and Trivedi (2005) discuss the large set of instruments that are potentially available in dynamic panel models and a series of papers provide information on efficient estimation strategies for these models (Arellano & Bond, 1998; Blundell & Bond, 2000; Blundell, Bond et al., 2000; Bond, 2002). Fortunately, our data set includes lagged measures of various dimensions of urbanicity and real price of consumption items that can be used to help provide identification, which will be discussed later.

We estimate robust standard errors using the two-step version of the Arellano-Bond system estimator (the one-step version uses a weighted matrix that does not depend on estimated parameters, while the two-step estimator may result in efficiency gains although the asymptotic distribution approximations may be less reliable due to the dependence of the two-step weighted matrix on estimated parameters) with a finite-sample correction (Windmeijer, 2005) using the -xtdpd- procedure (previously -xtabond2-) in Stata (Roodman, 2003). The consistency of the GMM estimator relies on the assumptions that there is no first-order serial autocorrelation in the errors of the level equation (2), and that the instrument matrix is truly exogenous and therefore valid to define the moment conditions. We thus perform two specification tests. First, we test for the presence of second-order autocorrelation in the differenced equation, which reflect whether the errors from the levels equations are serially uncorrelated (note that first-order autocorrelation in the differenced equation (3) is expected and does not signify an improper model specification). Second, we test for the exogeneity of the instruments using Sargan-Hansen's *J*-test, which is robust to heteroskedasticity and autocorrelation, and is asymptotically distributed as χ^2 in the number of restrictions.

This dynamic panel approach has been used in studies where both autocorrelation and endogeneity are of potential concern, including financial and investment economics (Carstensen & Toubal, 2004; Horioka & Wan, 2006), environmental economics (Arbués, Barberán et al., 2004), health care organization (Brown, Coffman et al., 2006; Mark, Harless et al., 2004) and the health-wealth relationship (Michaud & van Soest, 2008). For the empirical question of weight over time, the system GMM dynamic panel approach is ideal. This is the first paper to our knowledge to use it because it requires at least two consecutive waves of panel data (depending on the exact specification) and a large number of observations in each wave. We have six waves of over 4,000 unique men, with over 1,300 observations per wave. The results are also straightforward to interpret (in the same manner as with regression results).

We compare our two-step system GMM estimator to simple random effects (*exogenous regressor*) model that does not control for the correlation between the explanatory variables and the disturbance terms to see how the results differ. We expect the results of this estimation to be badly biased for the reasons laid out above[1].

IV. Data

This paper used comprehensive longitudinal data from the six most recent waves (1991, 1993, 1997, 2000, 2004 and 2006) of the China Health and Nutrition Survey (CHNS) on male adults (18 to 55 years old) interviewed during any of the survey waves. The CHNS were conducted in nine diverse provinces (Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, and Shandong) of China, and contains detailed individual-level information on income, diet, health and demography for all members of sampled households as well as detailed community level data on infrastructure, public services and facilities. A multistage, random cluster process was used to draw the sample surveyed in each of the provinces. Counties in the nine provinces were stratified by income and a weighted sampling scheme was used to randomly select four counties in each province. Villages and townships within the counties and urban and suburban neighborhoods within the cities were selected randomly into primary sampling units (PSUs). The same households were surveyed over time as best possible and newly formed households began to be surveyed in 1993.

Our analysis only looks at adult men for a number of reasons. First, while gender disparities would be interesting to uncover, the CHNS data did not contain unmeasured predictors such as metabolic rate that are more operative for women. Second, women often have the triple burden of work, children and domestic chores, which have competing effect on diet and

[1] As checks we also apply an *instrumental variables* estimator to equation (2) and an *instrumental variables estimator with fixed effects* (IV-FE) without $W_{i,t-1}$ in the model to control for both endogeniety and individual unobservable factors. We expect these two IV and IV-FE estimators to provide consistent parameter estimates but these estimators should be less efficient than the two-step system GMM estimator. We do not report the results but they are available upon request.

physical activity choices and limits the variance observed in data. Third, there is better variance in the data for adult men because men are more likely to experience occupational change, and have greater access to technologies that affect physical activity levels (Bell, Ge et al., 2002). After we limited the data to men between 18 and 55 years old and who were not disabled during a particular wave there were 16,883 person-wave observations made up by 10,935 men.

--- Table 1 about here ---

Of these, only 4,120 men had at least two consecutive waves of data, making up 8,645 observations (Table 1). Some of the loss of observations was due to the fact that those whose first survey was conducted in 2006 were not included in the analytic sample (850 observations made up of 643 men). In addition, Liaoning province was dropped from the survey and replaced by Heilongjiang province in 1997 (Heilongjiang was kept in henceforth) due to flooding in Liaoning that year. This meant that observations from adult men in Liaoning who were first collected in 1993 would not have made it to the analytic sample due to the missing data for 1997. Also, the 1991 and 1993 Heilongjiang sample did not exist.

Beyond these factors, there was also loss to follow up. To test whether attrition was systematic, we ran a Heckman selection model (Heckman, 1979). This two-stage estimation was based on whether an individual had two or more consecutive waves of data using observed exogenous characteristics (community urbanicity measures, prices, province, time, age, marital status, education attainment and predicted household income) in their first wave. The second stage was only conducted among those with two or more consecutive waves of data, and modeled the last observed weight of individuals using exogenous characteristics from both the first and the last observed waves. Results from the first step of the Heckman selection model suggests that the men who are younger, single, from Guangxi province (in the South), and who are from communities that generally scored lower on the various urbanicity measures in their first wave are more likely to be dropped from the analytic sample. However, the Wald test of independence in the errors between these two stages produced a χ^2-statistic of 1.63, meaning that we cannot reject the null that there is no correlation between the errors of these two equations (i.e., selection is not a problem). We also ran a Hausman specification test (Hausman, 1978) between the coefficients from the second equation and from basic OLS and found that we could not reject the null hypothesis that the difference in coefficients are not systematic ($\chi^2(35) = 1.43$). Therefore, selection does not appear to be a problem empirically.

Another potential problem with the data is the different interval lengths between each wave (varies from two to four years). We conducted sensitivity analyses by using interpolating data for 1994 and 2003 which ensured that the data would have a consistent three-year interval. A significant number of observations were lost as a result of the interpolation, but a comparison of the estimates from this partially interpolated data to those from the actual data among those who remained in the partially interpolated data showed that the results were not statistically different. Therefore, it appears that the different interval length between the waves is not a major concern.

A. *Dependent Variable*

Anthropometric data was collected by trained health workers during a comprehensive physical exam at a local clinic or at the respondent's home. Figure 1 shows that from 1991 to 2006, both weight and body mass index (BMI = weight in kg/(height in m)^2) rose significantly among adult men in China. In this analysis, we used weight as the dependent variable (but control for height) because biological and epidemiological studies have found that weight gain is mostly gained in the form of fat (rather than muscle mass or fat-free mass) among adults. This is particularly the case for populations that were previously undernourished or experienced weight fluctuations, either in childhood or adulthood (Dulloo, 2008; Dulloo, Jacquet et al., 2006; Remacle, Bieswal et al., 2004), and that concurrently have lowered their physical activity levels, as is the case in China. Because fat accumulation is well known to be highly associated with morbidity and mortality from cardiovascular diseases, type II diabetes, hypertension, and other nutrition-related non-communicable diseases (Folsom, Li et al., 1994; Matsuzawa, Nakamura et al., 1995; Nakamura, Tokunaga et al., 1994; Raymond, Leeder et al., 2006), weight and weight gain are important outcomes. Moreover, it is easier to interpret the results in terms of weight, because height does not change much within an adult population. As a check, we also ran the two-step system GMM with height as an endogenous variable.

--- Figure 1 about here ---

B. ***Key Explanatory Variables***

The key explanatory variables were lagged total physical activities, total caloric intake, proportion of energy from fat, energy from animal source foods, energy from fruits & vegetables, energy from edible oils and fats, energy from whole grain sources, energy from rice, drinking status and smoking status. Total physical activity was based on self- reported information on activity levels and time spent for up to two occupations, and time spent on four types of domestic activities (buying food, preparing food, doing laundry and childcare). These were combined with specific metabolic equivalent (MET) values based on the Compendium of Physical Activities (Ainsworth, Haskell et al., 2000) to derive MET-hours per week to account for both intensity of activities and time spent on activities. A unit of MET, is defined as the ratio of a person's working metabolic rate relative to his/her resting (basal) metabolic rate. There is additional information about leisure activities and travel activities, which was only available in the last four waves of the CHNS. However, limiting the analysis to only these last four waves would have severely compromised sample sizes, so only activities from occupations and domestic chores were included. Moreover, among men in China, these two domains made up the bulk of physical activities based on 1997-2006 data. Additional information on the creation of the physical activity measures are in a recent paper by Ng and colleagues (2009). In our analytic sample, physical activity levels among Chinese men fell significantly by 37 percentage points in a span of 15 years (Figure 2a).

Detailed consumption data at both the household and individual level were collected over three consecutive days (start day was randomly allocated from Monday to Sunday) in order to determine average daily dietary intake for each individual. Household food consumption was determined by examining changes in inventory from the beginning to the end of each day to account for food purchases, consumption and waste. Individual dietary intake for the same three consecutive days was surveyed for all individuals from 1991 onwards based on daily self-reported 24-hour recalls on all food consumed away-from-home and at-home. The collection of both household and individual dietary intake allowed for quality checks. Where significant discrepancies were found, the household and the individual in question were revisited and asked about their food consumption to resolve them (Wang, Ge et al., 2000).

The 1991 Food Composition Table (FCT) for China was utilized to calculate macronutrient intake values for the dietary data of 2000 and previous years (Institute of Nutrition and Food Hygiene, 1991). A new 2000 version of the FCT (Institute of Nutrition and Food Hygiene, 2002) was used for 2004 and 2006 surveys, and updated for new foods each year.

--- Figures 2a, 2b and 2c about here ---

We decided to use measures of total caloric intake (in kcals), proportion of energy from dietary fat (%) and from carbohydrates (%) as one set of diet variables. Another set of dietary variables that used food sources were also used, and included: total caloric intake, energy from animal source foods (%), energy from edible fats (%), energy from fruits and vegetables (%), energy from whole grain sources (%) and energy from rice (%). Additional dimensions on dietary intake will allow us to better capture the role of the various macronutrient as well as food sources in explaining weight. From the consumption data, we were able to get individual level measures of caloric fat, carbohydrate and protein, as well as calories from animal source foods, edible fats, fruits and vegetables and rice to allow us to determine these measures. Figure 2b shows that over the 15-year period, total caloric intake fell by about 17%, but the proportion of energy from dietary fat rose by six percentage points. Meanwhile, Table 2 and Figure 2c shows that the proportion of energy from animal source foods and edible oils and fats rose by 3.6 and 6.3 percentage points respectively, while energy from rice and energy from grains fell by 9.5 and 3.2 percentage points respectively.

The fact that both physical activity levels and total caloric intake have fallen while weight has increased suggest that declines in physical activity were more important than reductions in caloric intake. However, given that the proportion of energy from fat and carbohydrates has increased, as have energy from animal source foods, and oils and fats, it is likely that dietary composition may have a role in explaining the weight gain.

Dummy variables for being a drinker or smoker were included in the analyses because alcohol is calorically dense (7 kcals/gram) and so may result higher weights (Gordon & Doyle, 1986; Wannamethee & Shaper, 2003). Clinical and epidemiological studies have shown that smoking is consistently negatively related to body weight (Klesges, Meyers et al., 1989), possibly because nicotine increases energy expenditure and could reduce appetite (Hofstetter, Schutz et al., 1986; Williamson, Madans et al., 1991). The CHNS has information about whether respondents drank any beer or other alcohol beverage in the past year; and if they are current smokers. However, this is the extent of the smoking and drinking data available, which limits the variation since these are unlikely to change much from year to year. In this sample, the smoking and drinking prevalence among Chinese men declined with age (Table 2), likely due to the both the aging effect and mortality effect (smokers and drinkers might have higher mortality rates than non-smokers and non-drinkers).

--- Table 2 about here ---

Due to the longitudinal nature of this data, it is also important to account for aging and time or cohort effects. It is well known that basal metabolism changes with age (Tzankoff & Norris, 1978), and this change is likely non-linear. Thus, we used age group dummies in our models. We also consider the possibility that age and physically activity combined, or age and diet combined might affect weight differently and so we tested for the interaction of age with these key endogenous dependent variables. The biological literature suggests that the effects of physical activity and diet on weight for the age range in our empirical work (18 to 55) is unlikely to vary much by age (Henry, 2005; Keys, Taylor et al., 1973; Webb, 1981), but we felt that it would be useful to subject this supposition to an empirical test. Note also, that the time trend and cohort effects are difficult to separate out from the aging effects given the longitudinal nature of this data. While it is possible in theory to create cohorts for our analyses, the requirement of having consecutive waves of data severely limits the sample sizes for each cohort and compromises the system GMM estimation. Therefore, we have chosen to include both time dummies and age dummies as controls, but caution that we cannot provide an interpretation of the true age effect.

Other controls include marital status, living situation, highest education attainment, and predicted real household income tertile (created by using assets, occupations, education, age and household size to extrapolate for household with missing income data, and adjusting for community-specific CPI).

C. *Potential Instruments*

Potential instruments are time-varying and arguably exogenous dimensions of urbanicity and prices of food items for each community. We conducted specification tests on various sets of these variables to determine the final set of instruments for use in both the instrumental variables estimation and the dynamic panel estimation.

We used ten community-level measures of various dimensions of urbanicity: population, density, market accessibility, economic wellbeing, transportation, communications, education attainment, health facilities, sanitation and housing infrastructures. These reflect changes in the various dimensions of urbanicity over time and reflect the environment in which people function. Each of these dimensions was given a score from zero to ten and was comprised of data collected from local area administrators or official records. Ng (2009) explains in detail how these scores were created and how their distributions have changed over time.

The CHNS community-level measures of urbanicity have also been previously used in papers by Monda et al. (2007), and Zimmer et al. (2007). Figure 3 shows that over time, the communities on average had improvements in these dimensions. Urbanicity was not uniform across communities, with some communities experiencing declines in certain dimensions even though in general the average community might have seen improvements. China's household registration (*hukou*) system and the longitudinal nature of the CHNS data ensure that selection into communities and inclusion in the data was as independent of individual or household choices and behavior as possible.

--- Figure 3 about here ---

Prices may affect weight via consumption of various types of food items, alcohol and smoking. We included prices of food items that may be particularly important in the context of China, such as rice, flour, pork and oil, and prices of local beer and cigarettes. Community price surveys conducted on a set of sample stores and markets were used to provide price data. There were three sources of price information for a representative basket of goods. These include state store prices, free market prices collected from visits to stores in the communities surveyed, and authority price records published by the State Statistical Bureau (SSB) of China, which provides the provincial average. The state store prices were no longer used after the 1991–1992 price reform in China. Therefore, in almost all situations, the free market prices will be used as the basis, except when the goods studied were not sold in the free market, in which case, prices from the state stores will be used, followed by SSB recorded prices if the other two sources do not have the information. Farmers both produce and consume food, which adds complexity to the price issue. However, we would argue that the free market prices for the food can be seen as the opportunity cost of consuming instead of selling the produce. Hence, using free market prices (when available) is appropriate. Moreover, our models account for time-invariant individual effects, so as long as farming status does not change, we have handled some of the potential heterogeneity as it relates with individuals that also product their own food.

--- Figure 4 about here ---

Variations in prices across communities are due to both supply and demand side factors. On the supply side, agricultural production, transportation, marketing and distribution costs, imports of specific foods and other items, and availability of substitutes and complements can affect prices across communities. On the demand side, aggregate preferences or food fads may vary by communities. Most price changes in China are driven by supply side factors and exogenous economic decisions made at the provincial level by price commissions and other macroeconomic government decisions. In addition, while aggregate demand might affect prices, an individual's demand does not. Hence, the community prices can be considered exogenous variables that vary greatly over time and across communities as shown in the lack of a clear pattern in price changes among the ten most populous communities in the CHNS in Figure 4. In addition, there are also variations in inflation, measured by the Consumer Price Indices (CPI) across communities. A community-specific non-food CPI was derived by using a consumer basket of non-food items and the SSB's annual province and urban-rural specific consumer price index ratio because there is no published absolute non-food CPI for China that provides a way to compare provinces or urban and rural areas. Price and income variables were deflated by this constructed CPI with urban Liaoning province for 2006 equal to 1.00 with all other prices relative to this (CPC, 2006).

D. *Testing the validity of instruments*

We tested the null hypothesis that there is exogeneity by conducting a Hausman test between the model assuming the exogeneity of all explanatory variables and the instrumental variables model. The Hausman test showed that we can reject the null hypothesis that there is exogeneity ($\chi^2(21) = 52.76$). This suggests that we should use instrumental variables in the estimation (Hausman, 1978).

We first tested whether the instruments are correlated with the endogenous variables. The first-stage regression results use lagged values of the instruments and other control variables as explanatory variables for the endogenous variables of lagged weight, physical activity, dietary intake, smoking and drinking (see Table 3). We found that all but two of the lagged community urbanicity measures were statistically predictive of physical activity in the prior wave. In particular, the scores for educational institutions, sanitation, economic wellbeing and housing infrastructures were highly associated with declines in physical activity. The community urbanicity variables were also highly predictive of most of the dietary intake outcomes.

--- Table 3 about here ---

The lagged community price variables were most predictive of lagged weight, energy from fat, and energy from oils and fats. There are a few interesting results of note. For example, two key drivers of diet in China were particularly sensitive to price changes. A one percentage point increase in the price of pork appears to be related to a 1.83 percentage point decrease in energy from fat, as well as decreases in energy from animal source foods, and oils and fats. Also, a one percentage point increase in the price of oil appears to be related to a 1 percentage point decrease in energy from fat, as well as decreases in energy from grains, but increases in energy from rice. For the endogenous variables of lagged weight, energy from fat, carbohydrates, animal source foods, oils and fats, whole grain sources and rice, it appears that the ten community urbanicity measures and eight price variables satisfied the requirement that these instruments are correlated with them. For lagged physical activity and total caloric intake, the urbanicity measures were jointly significant. In general, the models did not perform as well for the health behaviors, which are not too surprising because we only had dichotomous variables for whether an individual smoke or drink in the past year or not. Because smoking and drinking can be addictive, it is unlikely that there is much variation in smoking and drinking status over time.

We then tested whether the instruments are uncorrelated to the error terms using a Hansen's *J*- tests of over-identification because there are more potential instruments then there are endogenous variables. The *J*-statistic follows a χ^2 distribution with degrees of freedom equal to the number of over-identifying restrictions rather than the total number of moment conditions. A rejection of the null hypothesis can imply that the instruments do not satisfy the orthogonality conditions (either because they are not truly exogenous, or because they are being incorrectly excluded from the regression), or that the model specification is incorrect. In standard IV models, the Sargan statistic is calculated instead. The Sargan statistic is a special case of

Hansen's *J*-statistic, which uses an estimate of the error variance from the IV regression estimated with the full set of over-identifying restrictions, and will generate a consistent estimator of the error variance under the null of instrument validity. We found that the Sargan test of over-identification cannot be rejected, meaning that the set of instruments appear to satisfy the requirement that the instruments be independent from unobserved error (Table 4, columns 3-4).

V. Results

A. *Exogenous Regressors*

The results for when we assume that the regressors are exogenous are consistent for either sets of diet variables (Table 4, columns 1 and 2). These results suggest that height is positively related to weight (p<0.01), while prior weight is positively related to current weight (p<0.01). The age and time effects are also statistically significant, with older men being heavier, as expected. An increase of 10 MET-hours per week of physical activity in the prior wave is associated with a weight loss of 0.0015 to 0.016kg (p<0.01). However, the results for total caloric intake, energy from fat and carbohydrates (column 1), and energy from the various food sources (column 2) were all not statistically significant. Smoking appears to lower weight, but drinking had no significant effect.

As mentioned earlier, these results can be biased and inconsistent due to endogeneity. The coefficients for the endogenous variables of physical activity and diet were all very different from those in the system GMM model, and in the case of some of the dietary variables had unexpected signs (but where not statistically significant). Moreover, the coefficient estimate for lagged weight is higher in this model compared to the GMM system approach, suggesting that time invariant unobserved heterogeneity is also a problem.

--- Table 4 about here ---

B. *Two-step system GMM*

The dynamic panel two-step procedure combines in a system GMM, a regression in differences over time, and a regression in levels. Recall that the consistency of the GMM estimation relies on the lack of autocorrelation of the residuals and the validity of the instrumental variables. The -xtdpd- postestimation procedure in Stata (Roodman, 2003) performs validation tests for these. In our estimation, the rejection of the presence of a second-order autocorrelation (i.e., the AR(2) *z*-statistic is not significant) satisfies the first criterion, and the rejection of the *J*-test of over-identification satisfies the second criterion (Table 4, columns 3-4).

The coefficient estimates from the GMM system dynamic panel estimation are interpreted as in a standard linear model. We found that, unsurprisingly, height is positively related to weight ($p<0.01$), while prior weight is positively related to current weight with a coefficient of 0.114 to 0.118 ($p<0.01$), a finding that is a little smaller to that found by Goldman and colleagues who also used a system GMM model to look at the effects of prices on BMI (Goldman, Lakdawalla et al., 2009). An increase of 10 MET-hours per week of physical activity in the prior wave is associated with a weight loss of 0.031kg ($p<0.01$), a larger coefficient than what was found in the results with no correction for endogeneity. Also, a one percentage point increase in energy from dietary fat in the prior wave was associated with a 1.74 kg weight gain, *ceteris paribus*, while a one percentage point increase in energy from carbohydrates in the prior wave was associated with a 1.25 kg weight gain, *ceteris paribus* (Column 3). As for energy from various food sources, energy from oils and fats (e.g., lard, butter, margarine, etc.) was the only dietary intake variable that was significant in affecting weight (Column 4). When height was considered an endogenous variable in the two-step system GMM model, the results were virtually the same (results available upon request).

We can also tell from these coefficients and the noted change in physical activity levels and dietary intake over the 1991 and 2006 period, how each of the main factors of interest may have contributed to weight gain. Table 5 provides examples of two hypothetical adult men. Example 5.1 is for a male who in 1991 was 30 years old with secondary school education, stayed married, lives with family from a medium income household and who is a persistent smoker and drinker. This man will naturally age over time and the time variables will change accordingly. By manipulating the various endogenous variables over time (either by keeping them the same as baseline throughout) or by allowing them to change based on the average values), we can see what proportion of weight gain might be due to physical activity, diet or aging and time changes. Example 5.2 is a male who entered the analysis in 1991 at age 40 with a primary school education, stayed married, lives with family from a low income household and who is a persistent smoker, but does not drink. For both of these examples, about 6.6 percent of weight change over the 15 year period was due to physical activity declines, while 3.3 to 3.5 percent was due to changes in diet.

--- Table 5 about here ---

Another way to interpret the results is by thinking about the short-run and long-run effects of a one time change in either physical activity or diet. Table 6 shows that for the average Chinese adult male, physical activity declines explain 6.1% of the weight gain over this 15 years period, while diet changes explain 2.9% to 3.8% (table 6, column 3). Column 6 of Table 6 show these effects based on the average change in physical activity and dietary intake per year.
We found at a one-time 9.5 MET-hour/week decrease in physical activity in a year would have a short-run (by next year) effect of a 0.029 kg weight gain but a long-run effect of 0.033 kg weight gain (or 6.9%). Meanwhile, a one-time 0.4 percentage point increase in energy from fat would have a short-run effect of a 0.06 kg weight gain and a long-run effect of a 0.08 kg weight gain. Moreover, a one-time 0.45 percentage point decrease in energy from carbohydrates would have a short-run effect of a 0.056 kg weight gain and a long-run effect of a 0.064 kg weight gain. These together can result in a long-run or persistent effect of 3.2% weight gain.

These may seem to be low values and rather discouraging findings for those trying to maintain or lose weight. However, a recent paper in the *Journal of the American Medical Association* (JAMA) by Katan and Ludwig (2010) that discusses the physiology of weight gain and loss lends some support to our finding. The authors surmise that "small changes in lifestyle would have a minor effect on obesity prevention". They explain that this is because any single change in diet or physical activity, even if permanent will elicit compensatory mechanisms that limit long-term effects on body weight. We find here in our estimation, that the modest coefficient for lagged weight (α) is in a sense limiting that the long-term effect.

VI. Conclusions

To our knowledge, this is one of the first papers that uses dynamic panel system GMM estimation to model the relationships between macro-level factors and micro-level behavior and their influence on weight. It is critical to analyze longitudinal data on dietary intake, physical activity, and other health behaviors over time to understand the dynamics of weight. Hence, the

GMM system dynamic panel approach is ideal for this research; it is preferred over typical reduced-form and IV models. We found that declines in physical activity and changes in the composition of people's diet are significantly related to weight among adult men in China. Of these two factors, the declines in physical activity seem to be a larger contributor to weight gain, although dietary intake (particularly dietary fat, carbohydrates and oils and fats) are also important.

Our findings are consistent with a review of studies by Schrawen and Westerterp (2007) that concluded that increased intake of dietary fat and a decreasing physical activity level are the most important environmental factors explaining the increased prevalence of obesity in westernized societies. Physical activity has been found to be a critical factor in body weight regulation in lean and obese individuals due to its protective role over time through both direct energy expenditure, improved physical fitness and resultant metabolic effects on lipid mobilization and oxidation and biochemical changes in the muscle fiber that contribute to improved regulation of body weight (Saris, 1998). Previous work has also hinted that physical activity may be a more successful strategy than dietary approaches to weight loss and maintenance among men (King, Frey-Hewitt et al., 1989).

Given the limited long-term effects of one-time physical activity or dietary changes to control weight, effective public health approaches to prevent overweight and obesity will require more fundamental changes in the food supply and the socio-economic and built environment. Strategies to increase physical activity levels at the workplace (Bell, Ge et al., 2001) and designing built environments that are safe and conducive for such transit or exercise modes (Bell, Ge et al., 2002; Forsyth, Hearst et al., 2008; Nagel, Carlson et al., 2008) can help prevent weight gain, as can policies in the form of higher taxes on automobiles, lower entry fees to parks and government run health facilities. Additionally, disincentives for automobile ownership can discourage motorized transportation and help reduce air pollution and provide more pleasant environments for outdoor exercise. We do not, however, know if these are cost-effective programmatic and policy options in China for increasing physical activity that will work to reduce the prevalence of overweight and obesity. As for dietary intake, there is some controversy regarding whether it is fat itself that increases weight or, that it is the fact that fat per gram is more energy dense (Bray, Paeratakul et al., 2004; Willett, 1998). Our findings that the proportion of energy from fat and carbohydates were important, but total caloric intake was not, seems to suggest that the former may be the case in this particular population.

Our estimated effects were modest, and we were only able to account for around 10% of the weight gain observed in the 15 year period. Much of the rest of the weight gain is probably due to aging and time effects. Unfortunately, we are unable to separate out the aging and the time and or cohort effects. In theory this is possible if we were to create separate cohorts of adult men, but given the reduced sample sizes due to the requirement of consecutive waves, this was not feasible. It is also possible that even with detailed data at the level in which exists in the CHNS, measurement error still exists and makes empirical analysis very challenging.

Moreover, due to the data requirements, we were only able to apply this modeling strategy on a specific population—adult (18-55 year old) men in China, which limits its applicability to the general population. While this analysis did not directly estimate the interacted effect of physical activity and dietary intake, it does so implicitly in the dynamic panel system GMM estimation approach by including lagged weight with controls for endogeneity in the physical activity and diet variables that affect weight. Certainly, there are also additional endogenous determinants of both diet and physical activity, such as technological changes related to home assets like refrigerators, rice cookers, microwaves, vacuum cleaners and washing machines that can be examined in the future. Future work should also consider regarding joint decisions about time and energy allocation among household members, instead of just considering individuals alone.

APPENDIX

An individual's utility in current period, t, depends on food consumption, F, physical activity (A), other health behaviors (such as smoking), other consumption, C, and current weight, W. Utility U increases with consumption of food and other things, but is increasing in weight only if current weight is less than ideal weight, $\ddot{W}$. Otherwise, utility declines with weight. The marginal utility of eating decreases as weight exceeds ideal weight, because eating increases weight. The assumption is that there is an ideal weight, $\ddot{W}$, holding other consumption constant. In other words, $\ddot{W}$ is the weight that would be chosen if achieving one's preferred weight were costless. This subjective ideal weight may or may not correspond to the weight that maximizes health or longevity, although it is likely to be influenced by concern with these factors. But the ideal weight is not necessarily the preferred weight in the economic sense because it does not consider the full range of costs and benefits of achieving it. In other words, a person's rationally chosen weight is the one that makes him the happiest given the existing costs and benefits of food consumption, physical activity and other consumption.

Because this model focuses on weight, we conceptualize food consumption simply as caloric intake, including calories from alcohol. Two other behaviors affect weight — physical activity, A, and smoking, S. Both affects utility directly and indirectly (as determinants of weight). An individual's physical activity level depends on the level of development, D, where she lives, such that $A_t = A(D_t)$.

In this dynamic problem, weight, W, is the state variable. Weight is a capital stock that depreciates at rate $1-\delta$ (where δ can be thought of as basal metabolism). Weight increases through consumption of food and alcohol, and decreases with physical activity and smoking. Thus, an individual's weight at time t, depends on prior weight, food and alcohol consumption F, smoking S, and physical activity level A:

(A1) $$W_t = (1-\delta)W_{t-1} + g(F_{t-1}, S_{t-1}, A(D_{t-1})),$$

where $\delta < 1$ and g is continuous, concave, increasing in food or alcohol consumption, decreasing in physical activity level, and decreasing in smoking level[2] ($g_F \geq 0$, $g_S \leq 0$, and $g_A \leq 0$).

Over multiple time periods, an individual's value function (or lifetime indirect utility) depends on the current period's utility and the value function from future time periods:

[2] Systematic reviews of clinical and epidemiological studies have shown that smoking is consistently negatively related to body weight (Klesges, Meyers et al., 1989)

(A2) $$v(W_t) = \max_{F,C,W} \{U_t(F_t, A_t, S_t, C_t, W_t) + \beta v(W_{t+1})\},$$

where β is the discount factor.

Individuals are subject to a budget constraint each period: $p_F F + p_S S + p_c C \le I$, where p_F, p_S and p_c are the prices of food (including alcohol), cigarettes, and other consumption goods respectively, and I is income. Consistent with existing literature, this does not account for borrowings and savings over time. Standardizing by p_c, the budget constraint as:

(A3) $$C \le I - p_F F - p_S S.$$

Combining Eq (A2) and Eq (A3), and taking the first order conditions with respect to F_t and C_t, and setting them to zero so that one is maximizing their utility, gives:

(A4)

$$U_F(F_t, A_t, S_t, (I_t - p_F F_t - p_S S), W_t) + \beta v'(W_{t+1}) g_F = p_F U_C[F_t, A_t, S_t, (I_t - p_F F_t - p_s S), W_t]$$

That is: Marginal utility of eating and drinking plus discounted marginal utility of weight in future period due to eating equals the marginal utility of consuming other goods. Similarly,

(A5) $U_A(F_t, A_t, S_t, (I_t - p_F F_t - p_S S), W_t) + \beta v'(W_{t+1}) g_A = 0$

(A6)

$$U_S(F_t, A_t, S_t, (I_t - p_F F_t - p_S S), W_t) + \beta v'(W_{t+1}) g_S = p_S U_C[F_t, A_t, S_t, (I_t - p_F F_t - p_s S), W_t]$$

Taking first order conditions of Eq (A2) with respect to W, we can get the envelope theorem:

(A7) $$v'(W_t) = U_W[F_t, A_t, S_t, (I_t - p_F F - p_S S), W_t] + \beta(1-\delta) v'(W_{t+1}),$$

which shows that the long run marginal value of weight is equal to the marginal utility of weight in the current period plus the discounted marginal utility of weight.

These will yield a steady-state in food and alcohol consumption, smoking, physical activity and weight as long as the marginal utility of food and alcohol consumption is falling in weight. Rewriting the optimality condition,

(A8) $$v'\{W_t + g[F_t, S_t, A(D_t)]\} = [(p_F U_C - U_F)/g_F] + [(p_S U_C - U_S)/g_S] + [U_A/g_A],$$

which is the marginal benefit of weight in the future equaling the marginal cost of spending on weight change via food, smoking or physical activity.

REFERENCES

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., et al. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc, 32*(9 Suppl), S498-504.

Alonso-Borrego, C., & Arellano, M. (1999). Symmetrically Normalized Instrumental-Variable Estimation Using Panel Data. *Journal of Business & Economic Statistics, 17*(1), 36-49.

Arbués, F., Barberán, R., & Villanúa, I. (2004). Price impact on urban residential water demand: A dynamic panel data approach. *Water Resrouce Research, 40*(W11402).

Arellano, M., & Bond, S. (1998). *Dynamic Panel Data Estimation Using DPD98 for Gauss: A Guide for Users*.Unpublished manuscript, London.

Astrup, A. (1999). Macronutrient balances and obesity: the role of diet and physical activity. *Public Health Nutr, 2*, 341-347.

Bell, A. C., Ge, K., & Popkin, B. M. (2001). Weight gain and its predictors in Chinese adults. *Int J Obes Relat Metab Disord, 25*(7), 1079-1086.

Bell, A. C., Ge, K., & Popkin, B. M. (2002). The road to obesity or the path to prevention: motorized transportation and obesity in China. *Obes Res, 10*(4), 277-283.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*(1), 115-143.

Blundell, R., & Bond, S. (2000). GMM Estimation with persistent panel data: an application to production functions. *Econometric Reviews, 19*(3), 321 - 340.

Blundell, R., Bond, S., & Windmeijer, F. (2000). *Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator*: Institute for Fiscal Studies.

Bond, S. R. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese Economic Journal, 1*(2), 141.

Bray, G. A., Paeratakul, S., & Popkin, B. M. (2004). Dietary fat and obesity: a review of animal, clinical and epidemiological studies. *Physiol Behav, 83*(4), 549-555.

Bray, G. A., & Popkin, B. M. (1998). Dietary fat intake does affect obesity! *Am. J. Clinical Nutrition, 68*(6), 1157-1173.

Brien, S. E., Katzmarzyk, P. T., Craig, C. L., & Gauvin, L. (2007). Physical activity, cardiorespiratory fitness and body mass index as predictors of substantial weight gain and obesity: the Canadian physical activity longitudinal study. *Can J Public Health, 98*(2), 121-124.

Brown, T. T., Coffman, J. M., Quinn, B. C., Scheffler, R. M., & Schwalm, D. D. (2006). Do Physicians Always Flee from HMOs? New Results Using Dynamic Panel Estimation Methods. *Health Services Research, 41*(2), 357-373.

Cameron, A. C., & Trivedi, P., K. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.

Carstensen, K., & Toubal, F. (2004). Foreign direct investment in Central and Eastern European countries: a dynamic panel analysis. *Journal of Comparative Economics, 32*(1), 3-22.

Cawley, J. (2004). An economic framework for understanding physical activity and eating behaviors. *American Journal of Preventive Medicine, 27*(3), 117-125.

Chou, S. Y., Grossman, M., & Saffer, H. (2004). An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics, 23*(3), 565-587.

Cutler, D. M., Glaeser, E. L., & Shapiro, J. M. (2003). Why have Americans become more obese? *Journal of Economic Perspectives, 17*(3), 93-118.

Drewnowski, A., & Darmon, N. (2005). The economics of obesity: dietary energy density and energy cost. *American Journal of Clinical Nutrition, 82*(1), 265s-273s.

Dulloo, A. G. (2008). Thrifty energy metabolism in catch-up growth trajectories to insulin and leptin resistance. *Best Practice & Research Clinical Endocrinology & Metabolism, 22*(1), 155-171.

Dulloo, A. G., Jacquet, J., Seydoux, J., & Montani, J. P. (2006). The thrifty `catch-up fat' phenotype: its impact on insulin sensitivity during growth trajectories to obesity and metabolic syndrome. *Int J Obes, 30*(S4), S23-S35.

Folsom, A. R., Li, Y., Rao, X., Cen, R., Zhang, K., Liu, X., et al. (1994). Body mass, fat distribution and cardiovascular risk factors in a lean population of South China. *Journal of Clinical Epidemiology, 47*(2), 173-181.

Forsyth, A., Hearst, M., Oakes, J. M., & Schmitz, K. H. (2008). Design and Destinations: Factors Influencing Walking and Total Physical Activity. *Urban Stud, 45*(9), 1973-1996.

Frisancho, A. R. (2003). Reduced rate of fat oxidation: A metabolic pathway to obesity in the developing nations. *American Journal of Human Biology, 15*(4), 522-532.

Goldman, D., Lakdawalla, D., & Zheng, Y. (2009). *Food Prices and the Dynamics of Body Weight*.Unpublished manuscript.

Gordon, T., & Doyle, J. T. (1986). Alcohol Consumption and Its Relationship to Smoking, Weight, Blood Pressure, and Blood Lipids: The Albany Study. *Arch Intern Med, 146*(2), 262-265.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica, 46*(6), 1251-1271.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica, 47*(1), 153-161.

Henry, C. J. K. (2005). Basal metabolic rate studies in humans: measurement and development of new equations. *Public Health Nutrition, 8*(7a), 1133-1152.

Hofstetter, A., Schutz, Y., Jequier, E., & Wahren, J. (1986). Increased 24-hour energy expenditure in cigarette smokers. *N Engl J Med, 314*(2), 79-82.

Horioka, C. Y.-J., & Wan, J.-M. (2006). *The Determinants of Household Saving in China: A Dynamic Panel Analysis of Provincial Data*: National Bureau of Economic Research, Inc.

Institute of Nutrition and Food Hygiene. (1991). *Chinese Food Composition Tables 1991* Beijing, PRC: People's Health Publishing House

Institute of Nutrition and Food Hygiene. (2002). *China Food Composition Table 2002*. Beijing, PRC: Beijing Medical University Publishing House.

James, W. P., & Ralph, A. (1999). New understanding in obesity research. *Proc Nutr Soc, 58*(2), 385-393.

Katan, M. B., & Ludwig, D. S. (2010). Extra Calories Cause Weight Gain--But How Much? *JAMA, 303*(1), 65-66.

Keys, A., Taylor, H. L., & Grande, F. (1973). Basal metabolism and age of adult man. *Metabolism, 22*(4), 579-587.

King, A. C., Frey-Hewitt, B., Dreon, D. M., & Wood, P. D. (1989). Diet vs exercise in weight maintenance. The effects of minimal intervention strategies on long-term outcomes in men. *Arch Intern Med, 149*(12), 2741-2746.

Klesges, R. C., Meyers, A. W., Klesges, L. M., & LaVasque, M. E. (1989). Smoking, body weight, and their effects on smoking behavior: A comprehensive review of the literature. *Psychological Bulletin, 106*(2), 204-230.

Kuchler, F., Tegene, A., Harris, J. M., & United States. Dept. of Agriculture. Economic Research Service. (2004). *Taxing snack foods what to expect for diet and tax revenues*. Washington, D.C.: Economic Research Serivce U.S. Dept. of Agriculture.

Lakdawalla, D., & Philipson, T. (2009). The growth of obesity and technological change. *Economics & Human Biology, 7*(3), 283-293.

Lakdawalla, D., Philipson, T. J., & Bhattacharya, J. (2005). Welfare-enhancing technological change and the growth of obesity. *American Economic Review, 95*(2), 253-257.

Mark, B. A., Harless, D. W., McCue, M., & Xu, Y. (2004). A Longitudinal Examination of Hospital Registered Nurse Staffing and Quality of Care. *Health Services Research, 39*(2), 279-300.

Matsuzawa, Y., Nakamura, T., Shimomura, I., & Kotani, K. (1995). Visceral fat accumulation and cardiovascular disease. *Obes Res, 3 Suppl 5*, 645S-647S.

Michaud, P.-C., & van Soest, A. (2008). Health and wealth of elderly couples: Causality tests using dynamic panel data models. *Journal of Health Economics, 27*(5), 1312-1325.

Miller, W. C., Lindeman, A. K., Wallace, J., & Niederpruem, M. (1990). Diet composition, energy intake, and exercise in relation to body fat in men and women. *Am. J. Clinical Nutrition, 52*(3), 426-430.

Monda, K. L., Gordon-Larsen, P., Stevens, J., & Popkin, B. M. (2007). China's transition: The effect of rapid urbanization on adult occupational physical activity. *Soc Sci Med, 64*(4), 858-870.

Moore, M. S. (2000). Interactions between physical activity and diet in the regulation of body weight. *Proc Nutr Soc, 59*(2), 193-198.

Nagel, C. L., Carlson, N. E., Bosworth, M., & Michael, Y. L. (2008). The Relation between Neighborhood Built Environment and Walking Activity among Older Adults. *Am. J. Epidemiol., 168*(4), 461-468.

Nakamura, T., Tokunaga, K., Shimomura, I., Nishida, M., Yoshida, S., Kotani, K., et al. (1994). Contribution of visceral fat accumulation to the development of coronary artery disease in non-obese men. *Atherosclerosis, 107*(2), 239-246.

Ng, S. W., Norton, E. C., & Popkin, B. M. (2009). Why have physical activity levels declined among Chinese adults? Findings from the 1991-2006 China health and nutrition surveys. *Soc Sci Med, 68*(7), 1305-1314.

Philipson, T. J., & Posner, R. A. (2003). The long-run growth in obesity as a function of technological change. *Perspect Biol Med, 46*(3 Suppl), S87-107.

Popkin, B. M., Kim, S., Rusev, E. R., Du, S., & Zizza, C. (2006). Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases. *Obesity Reviews, 7*(3), 271-293.

Pratt, M., Macera, C. A., Sallis, J. F., O'Donnell, M., & Frank, L. D. (2004). Economic interventions to promote physical activity: Application of the SLOTH model. *American Journal of Preventive Medicine, 27*(3, Supplement 1), 136-145.

Prentice, A. M., & Jebb, S. A. (2004). Energy Intake/Physical Activity Interactions in the Homeostasis of Body Weight Regulation. *Nutrition Reviews, 62*(s2), S98-S104.

Rashad, I. (2006). Structural estimation of caloric intake, exercise, smoking, and obesity. *The Quarterly Review of Economics and Finance, 46*(2), 268-283.

Rashad, I., & Grossman, M. (2004). The economics of obesity. *Public Interest*(156), 104-112.

Raymond, S. U., Leeder, S., & Greenberg, H. M. (2006). Obesity and cardiovascular disease in developing countries: a growing problem and an economic threat. *Current Opinion in Clinical Nutrition and Metabolic Care, 9*(2), 111-116.

Remacle, C., Bieswal, F., & Reusens, B. (2004). Programming of obesity and cardiovascular disease.(International Symposium "Childhood Obesity: From basic knowledge to effective prevention" and the 14th Workshop "European Childhood Obesity Group".)(Paper). *International Journal of Obesity, 28*(S3), S46.

Roodman, D. (2003). XTABOND2: Stata module to extend xtabond dynamic panel data estimator (Version S435901): Boston College Department of Economics.

Saris, W. H. (1998). Fit, fat and fat free: the metabolic aspects of weight control. *Int J Obes Relat Metab Disord, 22 Suppl 2*, S15-21.

Schrauwen, P., & Westerterp, K. R. (2007). The role of high-fat diets and physical activity in the regulation of body weight. *British Journal of Nutrition, 84*(04), 417-427.

Schroeter, C., Lusk, J., & Tyner, W. (2008). Determining the impact of food price and income changes on body weight. *Journal of Health Economics, 27*(1), 45-68.

Sturm, R. (2005). Economics and physical activity - A research agenda. *American Journal of Preventive Medicine, 28*(2), 141-149.

Taylor, C. B., Jatulis, D. E., Winkleby, M. A., Rockhill, B. J., & Kraemer, H. C. (1994). Effects of Life-Style on Body Mass Index Change. *Epidemiology, 5*(6), 599-603.

Tryon, W. W. (1987). Activity as a function of body weight. *Am. J. Clinical Nutrition, 46*(3), 451-455.

Tzankoff, S. P., & Norris, A. H. (1978). Longitudinal changes in basal metabolism in man. *J Appl Physiol, 45*(4), 536-539.

Wang, Y., Ge, K., & Popkin, B. M. (2000). Tracking of body mass index from childhood to adolescence: a 6-y follow-up study in China. *Am J Clin Nutr, 72*(4), 1018-1024.

Wannamethee, S. G., & Shaper, A. G. (2003). Alcohol, body weight, and weight gain in middle-aged men. *Am. J. Clinical Nutrition, 77*(5), 1312-1317.

Webb, P. (1981). Energy expenditure and fat-free mass in men and women. *Am J Clin Nutr, 34*(9), 1816-1826.

Willett, W. C. (1998). Is dietary fat a major determinant of body fat? *Am J Clin Nutr, 67*(3), 556S-562.

Williamson, D. F., Madans, J., Anda, R. F., Kleinman, J. C., Giovino, G. A., & Byers, T. (1991). Smoking cessation and severity of weight gain in a national cohort. *N Engl J Med, 324*(11), 739-745.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics, 126*(1), 25-51.

Zimmer, Z., Kaneda, T., & Spess, L. (2007). An examination of urban versus rural mortality in China using community and individual data. *J Gerontol B Psychol Sci Soc Sci, 62*(5), S349-357.

TABLE 1 — SAMPLE SIZE OF MEN FROM THE CHNS 1991-2006

Consecutive waves	Unit of observation	Wave 1991	1993	1997	2000	2004	2006	Total
T ≥ 2	Individuals		1,138	542	779	728	993	4,120
	Person-wave		2,014	1,381	1,756	1,743	1,751	8,645
T =1	Individuals	2,398	558	1,280	984	892	643	6,755
	Person-wave	2,873	654	1,487	1,314	1,060	850	8,238
Total	Individuals	2,398	1,696	1,822	1,763	1,620	1,636	10,935
	Person-wave	2,873	2,668	2,868	3,070	2,803	2,601	16,883

TABLE 2—WEIGHT, PHYSICAL ACTIVITY LEVELS, DIETARY INTAKE, AND OTHER DEMOGRAPHIC CHARACTERISTICS AMONG ADULT MEN IN CHINA (CHNS 1991-2006)

	Year						Change (1991-2006)
	1991	**1993**	**1997**	**2000**	**2004**	**2006**	
Weight (kg)	59.25 (8.46)	59.71 (8.49)	61.65 (9.56)	63.70 (10.28)	65.93 (13.72)	66.43 (12.22)	7.18 **
Height (cm)	166.1 (6.26)	165.95 (6.17)	166.60 (6.36)	167.31 (6.38)	167.66 (6.65)	167.88 (6.88)	1.78 **
BMI (kg/m^2)	21.43 (2.50)	21.65 (2.55)	22.15 (2.82)	22.71 (3.07)	23.40 (4.35)	23.54 (4.01)	2.11 **
Work & Domestic physical activity level (MET-hrs/week)	389.84 (220.37)	356.92 (217.01)	346.40 (215.97)	305.97 (202.97)	246.00 (180.18)	247.09 (177.90)	-142.75 **
Total caloric Intake (kcal)	2972.73 (826.07)	2872.65 (922.26)	2612.81 (717.89)	2618.09 (807.06)	2530.91 (804.02)	2458.11 (774.49)	-514.62 **
Energy from dietary fat (%)	21.72	23.06	25.50	28.22	28.03	27.79	6.07 **
Energy from carbohydrates (%)	66.69	65.06	62.71	60.05	60.09	59.9	-6.79 **
Energy from Animal Source Foods (%)	8.96	9.74	9.64	11.03	10.53	12.56	3.60 **
Energy from Oils and Fats (%)	8.18	9.24	12.02	13.43	14.54	14.52	6.34 **
Energy from Fruits and Vegetables (%)	2.22	2.42	1.82	1.85	2.15	1.99	-0.23
Energy from Rice (%)	50.03	50.22	44.36	43.61	42.51	40.56	-9.48 **
Energy from Grains (%)	26.62	24.21	26.83	24.44	24.16	23.46	-3.15 **
Smoker (%)	68.57	65.99	63.11	61.28	60.07	57.52	-11.05 **
Drinker (%)	67.77	64.14	67.12	65.27	64.15	63.32	-4.45 *
Age (year)	35.18 (9.93)	35.83 (10.08)	36.73 (10.19)	38.08 (10.10)	39.72 (9.95)	40.41 (9.78)	5.23 **
Married (%)	79.69	78.60	78.70	80.0	82.08	83.93	4.24 *
Live alone (%)	0.21	0.19	0.21	0.23	0.32	0.31	0.10 **
No education (%)	15.12	11.12	9.75	6.44	4.37	6.50	-8.62 **
Highest education is primary school (%)	62.68	65.61	64.66	62.79	61.75	54.94	-7.74 **
Highest education is secondary school (%)	15.26	16.50	17.46	18.49	19.75	20.57	5.31 **
Highest education is technical school (%)	3.19	3.98	4.12	5.77	7.62	9.00	5.81 **
Has university degree or higher (%)	3.75	2.79	4.01	6.51	6.51	8.99	5.24 **
Predicted household income (2006 yuan)	2228.2 (1443.17)	2261 (3746.15)	2545.71 (2362.72)	3573.33 (3432.76)	4431.98 (4419.80)	5243.65 (8047.30)	3015.45 **
Number of observations	2851	2649	2841	2980	2795	2601	

Notes: Standard Deviations in parentheses.
* difference between 1991 and 2006 is significant at 5%; ** significant at 1%

TABLE 3—URBANICITY AND PRICES ON WEIGHT, PHYSICAL ACTIVITY, DIETARY INTAKE, SMOKING AND DRINKING AMONG CHINESE MEN

Community Factors	**Weight (kg)**	**Activity (MET-hrs/wk)**	**Total caloric intake (kcal)**	**P(Smoke)**	**P(Drink)**
Population score (0-10)	-0.0282	0.229	6.806	0.00246	-0.00189
	(0.0550)	(1.509)	(7.746)	(0.00276)	(0.00294)
Density score (0-10)	0.00872	**-3.083** $^{+}$	-6.583	-0.00116	**0.00907 ****
	(0.0576)	(1.628)	(7.963)	(0.00330)	(0.00311)
Market Accessibility score (0-10)	-0.000103	**-2.974 ****	**-11.612 ***	**-0.00390 ***	**-0.00600 ****
	(0.0332)	(1.147)	(5.321)	(0.00196)	(0.00209)
Transportation score (0-10)	0.0261	**-3.387 ***	**-16.446 ***	0.000022	0.002
	(0.0465)	(1.499)	(8.120)	(0.00237)	(0.00253)
Communications score (0-10)	0.0802	-2.450	-5.658	-0.00176	0.00131
	(0.0549)	(1.960)	(9.0835)	(0.00283)	(0.00367)
Economy score (0-10)	0.0733	**-7.564 ****	11.401	0.00106	0.00115
	(0.0612)	(1.883)	(8.440)	(0.00331)	(0.00352)
Educational Institution score (0-10)	**0.171 ****	**-10.645 ****	**-35.515 ****	-0.0000166	0.00275
	(0.0583)	(2.271)	(10.454)	(0.00346)	(0.00410)
Health Facilities score (0-10)	0.0692	**-3.975 ***	-5.0896	0.00462	0.00242
	(0.0547)	(1.928)	(8.601)	(0.00315)	(0.00311)
Sanitation infrastructure score (0-10)	0.0483	**-8.997 ****	1.344	0.00273	-0.00111
	(0.0409)	(1.434)	(8.523)	(0.00216)	(0.00271)
Housing infrastructure score (0-10)	0.116 $^{+}$	**-6.654 ****	-5.956	-0.00126	-0.00422
	(0.0678)	(2.279)	(11.940)	(0.00229)	(0.00398)
Log real price of Rice (yuan/kg)	**0.700**	-0.294	-14.630	**0.0439 +**	-0.0111
	(0.448)	(14.0969)	(89.795)	(0.0253)	(0.0344)
Log real price of Flour (yuan/kg)	**-1.596 ****	-3.147	-12.161	-0.00195	-0.0370
	(0.384)	(15.673)	(74.461)	(0.0264)	(0.0313)
Log real price of Pork (yuan/ kg)	**-2.538 ****	-10.207	**162.713 ****	**0.0683 ****	-0.00110
	(0.477)	(11.495)	(75.777)	(0.0258)	(0.0261)
Log real price of Chicken (yuan/kg)	**-1.344 ****	17.0144	1.604	0.0137	-0.00445
	(0.399)	(11.179)	(65.275)	(0.0262)	(0.0227)
Log real price of Oil (yuan/liter)	**-0.485 ***	3.336	59.329	0.00118	-0.0148
	(0.252)	(8.225)	(41.066)	(0.0142)	(0.0149)
Log real price of Beer (yuan/bottle)	-0.420	-6.605	-14.732	-0.00622	-0.0330
	(0.460)	(11.775)	(69.814)	(0.0265)	(0.0288)
Log real price of Cigarettes (yuan/box)	0.073	-12.904	-33.568	**-0.0195 +**	-0.00706
	(0.211)	(8.0373)	(37.0418)	(0.0119)	(0.0137)
Non-Food Consumer Price Index (100=urban Liaoning)	**-8.258 ****	22.376	63.164	**0.207 ***	0.0494
	(2.209)	(62.613)	(260.242)	(0.0979)	(0.117)
Observations	8645	8645	8645	8645	8645
Number of Individuals	4120	4120	4120	4120	4120
Overall Statistic $\chi^2(32)$	1580.76 **	2264.12 **	173.17 **	309.91 **	337.66 **
Joint test of significance: all community variables $\chi^2(18)$	174.86 **	666.81 **	53.49 **	35.87 **	33.44 *
Joint test of significance: urbanicity measures $\chi^2(10)$	51.60 **	445.39 **	37.95 **	11.25	21.68 *
Joint test of significance: price variables $\chi^2(8)$	107.21 **	6.15	9.06	20.65 **	7.11

Notes: Controlling for height, time, age, marital status, living situation, education, and predicted household income. Robust standard errors in parentheses.
+ significant at 10%;* significant at 5%; ** significant at 1%

TABLE 3, CONTINUED—URBANICITY AND PRICES ON WEIGHT, PHYSICAL ACTIVITY, DIETARY INTAKE, SMOKING AND DRINKING AMONG CHINESE MEN

Community Factors	% energy from						
	Fat	**Carbohy.**	**Animal Source foods**	**Oils and fats**	**Fruits and vegetables**	**Whole grain sources**	**Rice**
Population score (0-10)	-0.0603	**0.0589 ***	0.000539	-0.000959	**0.000534** $^{+}$	**-0.00560** $^{+}$	0.00533
	(0.0807)	(0.0257)	(0.000628)	(0.000935)	(0.000320)	(0.00343)	(0.00344)
Density score (0-10)	**0.132** $^{+}$	**-0.307 ****	**0.00215 ****	-0.000291	0.0000969	0.00290	-0.00657 $^{+}$
	(0.0785)	(0.0891)	(0.000581)	(0.00108)	(0.000195)	(0.00368)	(0.00382)
Market Accessibility score (0-10)	-0.0362	0.107	0.0000891	-0.000161	-0.0000549	-0.00009	-0.0000449
	(0.0670)	(0.0721)	(0.000462)	(0.000639)	(0.000162)	(0.00143)	(0.00170)
Transportation score (0-10)	**0.142** $^{+}$	**-0.204 ***	**0.00143 ***	0.000707	0.000129	**-0.00423 ***	0.00249
	(0.0858)	(0.0961)	(0.000658)	(0.000932)	(0.000175)	(0.00199)	(0.00204)
Communications score (0-10)	**0.269 ***	**-0.282 ****	0.000218	0.00187	0.000124	0.00259	**-0.00489 ***
	(0.115)	(0.110)	(0.000858)	(0.00119)	(0.000377)	(0.00248)	(0.00257)
Economy score (0-10)	**0.188** $^{+}$	-0.222	**0.00173 ***	0.00162	-0.000457	**0.00653 ****	**-0.00994 ****
	(0.113)	(0.121)	(0.000859)	(0.00121)	(0.000432)	(0.00216)	(0.00252)
Educational Institution score (0-10)	**0.333 ****	**-0.332 ****	0.00122	0.000964	-0..000454	-0.00396	-0.000268
	(0.112)	(0.127)	(0.000801)	(0.00126)	(0.000450)	(0.00412)	(0.00415)
Health Facilities score (0-10)	**0.203 ***	**0.293 ****	**0.00181 ***	0.000260	-0.0000448	-0.00300	-0.000788
	(0.088)	(0.0851)	(0.000788)	(0.000996)	(0.000213)	(0.00216)	(0.00246)
Sanitation infrastructure score (0-10)	**0.198 ****	-0.159	0.00113	-0.00101	**-0.000593** $^{+}$	-0.000477	0.000761
	(0.0769)	(0.088)	(0.000706)	(0.000866)	(0.000358)	(0.00224)	(0.00225)
Housing infrastructure score (0-10)	**0.526 ****	**-0.600 ****	**0.00470 ****	0.00168	-0.000174	**-0.0146 ****	**0.00858 ***
	(0.123)	(0.133)	(0.000981)	(0.00127)	(0.000358)	(0.00319)	(0.00370)
Log real price of Rice (yuan/kg)	-0.901	1.104	**-0.0186 ****	**0.0207 ***	-0.00329	**0.153 ****	**-0.165 ****
	(0.938)	(0.955)	(0.00760)	(0.0100)	(0.00487)	(0.0299)	(0.0287)
Log real price of Flour (yuan/kg)	**1.672** $^{+}$	-0.237	0.0110	0.000746	**0.00681 ***	**-0.125 ****	**0.127 ****
	(0.878)	(0.908)	(0.00764)	(0.00814)	(0.00327)	(0.0209)	(0.0243)
Log real price of Pork (yuan/ kg)	**-1.832 ****	-1.520	**-0.0229 ****	**-0.0274 ****	0.000430	**-0.0872 ****	**0.0910 ****
	(0.724)	(0.806)	(0.00683)	(0.00888)	(0.00176)	(0.0169)	(0.0185)
Log real price of Chicken (yuan/kg)	**-1.376 ***	**1.396 ***	-0.00174	**-0.0229 ****	-0.00227	**-0.0319** $^{+}$	**0.0729 ****
	(0.650)	(0.695)	(0.00549)	(0.00837)	(0.00331)	(0.0179)	(0.0189)
Log real price of Oil (yuan/liter)	**-0.987 ***	-0.745	0.00147	**-0.00986** $^{+}$	-0.00116	**-0.0383 ****	**0.0499 ****
	(0.465)	(0.487)	(0.00335)	(0.00551)	(0.00206)	(0.0111)	(0.0125)
Log real price of Beer (yuan/bottle)	0.203	-1.203	**0.0255 ****	**-0.0319 ****	0.00173	**-0.0417** $^{+}$	**0.0594 ***
	(0.836)	(0.971)	(0.00729)	(0.00870)	(0.00226)	(0.0233)	(0.0249)
Log real price of Cigarettes (yuan/box)	**0.497**	-0.635	0.00456	**0.00833 +**	0.00148	**-0.0188 ***	0.00904
	(0.394)	(0.404)	(0.00292)	(0.00469)	(0.00118)	(0.0096)	(0.0112)
Non-Food Consumer Price Index (100=urban Liaoning)	**14.412 ****	**-18.103****	**0.162 ****	-0.0399	**0.0133** $^{+}$	**-0.367 ****	**0.226 ***
	(2.997)	(3.646)	(0.0264)	(0.0344)	(0.00781)	(0.0892)	(0.0954)
Observations	8645	8645	8642	8642	8642	8642	8642
Number of Individuals	4120	4120	4120	4120	4120	4120	4120
Overall Statistic $\chi^2(32)$	965.25 **	958.79**	829.50 **	416.32 **	80.85 **	220.20 **	487.89 **
Joint test of significance: all community variables $\chi^2(18)$	454.40 **	465.39 **	584.25 **	93.72 **	34.04 **	152.78 **	190.30 **
Joint test of significance: urbanicity measures $\chi^2(10)$	209.81 **	237.09 **	198.52 **	16.27 +	15.54	49.47 **	31.09 **
Joint test of significance: price variables $\chi^2(8)$	51.82 **	58.15 **	89.67 **	55.39 **	16.82 *	115.26 **	149.51 **

Notes: Controlling for height, time, age, marital status, living situation, education, and predicted household income. Robust standard errors in parentheses.
+ significant at 10%;* significant at 5%; ** significant at 1%

TABLE 4—RESULTS FROM ESTIMATING DETERMINANTS OF WEIGHT AMONG CHINESE MEN

Key Explanatory Variables	Exogenous regressors (1)	(2)	Two-step system GMM (3)	(4)
Current Height (cm)	**0.395 ****	**0.393 ****	**1.643 ****	**1.592 ****
	(0.0317)	(0.0314)	(0.295)	(0.190)
Age: 18-30	**-0.941 ****	**-0.920 ****	-3.468	**-4.308 ****
	(0.307)	(0.306)	(2.217)	(1.724)
Age: 30-45	0.321	**0.341 ***	-1.194	-0.106
	(0.169)	(0.169)	(1.366)	(1.0816)
Last wave was in 2000	**1.300 ***	1.0748	**0.597 ***	**0.478 ***
	(0.592)	(0.606)	(0.271)	(0.232)
Last wave was in 2004	**1.664 ****	**1.513 ****	0.265	0.0693
	(0.604)	(0.610)	(0.477)	(0.398)
Last wave was in 2006	0.923	0.718	-0.132	-0.309
	(0.681)	(0.693)	(0.613)	(0.516)
Lagged (t-1)				
Weight (kg)	**0.501 ****	**0.499 ****	**0.114 ****	**0.118 ****
	(0.041)	(0.0408)	(0.0425)	(0.0329)
Work & Domestic Physical activity level (METs-hours/week)	**-0.00158 ****	**-0.00154 ****	**-0.00307 ****	**-0.00307 ****
	(0.000495)	(0.000493)	(0.00125)	(0.000983)
Total Caloric Intake (kcal)	0.000170	0.000163	0.000146	0.0000762
	(0.00095)	(0.0000928)	(0.000157)	(0.000123)
Energy from Fat (%)	-0.00445		**0.174 ****	
	(0.0166)		(0.0426)	
Energy from Carbohydrates (%)	-0.0169		**0.125 ****	
	(0.0159)		(0.0428)	
Energy from Animal Source Foods (%)		-0.0230		0.0630
		(0.0221)		(0.0437)
Energy from Edible Oils & Fats (%)		0.00125		**0.0890 ***
		(0.0196)		(0.0411)
Energy from Fruits and Vegetables (%)		-0.0395		0.0178
		(0.0251)		(0.0711)
Energy from Whole grains (%)		-0.0103		0.0429
		(0.0187)		(0.0419)
Energy from Rice (%)		-0.0195		0.0400
		(0.0177)		(0.0408)
Smoker	**-0.524 ****	**-0.524 ****	0.559	**0.903 ***
	(0.188)	(0.188)	(0.521)	(0.431)
Drinker	0.262	0.294	0.713	0.457
	(0.215)	(0.214)	(0.394)	(0.353)
Instruments	None 18 Community variables included as controls in model		Differenced Eq: Lagged Z_i and Δ community variables Levels Eq: Lagged difference for Z_i for (t-1) and prior	
Number of Instruments used	None		75	99
Overall Statistic	$\chi^2(38)$= 6510.32**	$\chi^2(41)$= 6793.06**	$\chi^2(20)$= 247.93**	$\chi^2(23)$= 376.41**
Sargan/Hansen test of over-identification			$\chi^2(54)$= 79.62**	$\chi^2(75)$= 132.09**
Test for Autocorrelation				
AR(1) in first differences (*z*-statistic)			-3.595 **	-3.934 **
AR(2) in first differences (*z*-statistic)			-0.838	-0.422
N	8,645	8,642	8,645	8,642
Unique individuals	4,120	4,120	4,120	4,120

Notes: Controlling for marital status, living situation, education, and predicted household income. Robust standard errors in parentheses. * significant at 5%; ** significant at 1%

TABLE 5—CONTRIBUTION OF PHYSICAL ACTIVITY AND DIET ON WEIGHT GAIN: TWO EXAMPLES

Example 5.1: Male who entered analysis in 1991 at 30 years old with secondary school education, stayed married, lives with family from a medium income household and who is a persistent smoker and drinker	**Explained by model (3)**			**Explained by model (4)**		
	In kg	**As % of actual weight change ***	**As % of explained change**	**In kg**	**As % of actual weight change ***	**As % of explained change**
If only age and time changed	5.77	80.31	89.03	5.02	69.92	87.37
If age, time and average PA changed	6.24	86.91	96.34	5.49	76.53	95.63
If age, time and average diet changed	6.00	83.62	92.69	5.27	73.42	91.74
If age, time and average PA and diet changed	6.48	90.21	100	5.75	80.03	100
Contribution of change in PA on weight change	0.47	6.60	7.31	0.47	6.6	8.26
Contribution of changes in Diet on weight change	0.23	3.31	3.66	0.25	3.5	4.37
Example 5.2: Male who entered analysis in 1991 at 40 years old with primary school education, stayed married, lives with family from a low income household and who is a persistent smoker, but does not drink	**Explained by model (3)**			**Explained by model (4)**		
	In kg	**As % of actual weight change ***	**As % of explained change**	**In kg**	**As % of actual weight change ***	**As % of explained change**
If only age and time changed	3.94	54.91	84.72	2.34	32.53	76.31
If age, time and average PA changed	4.42	61.51	94.91	2.81	39.14	91.81
If age, time and average diet changed	4.18	58.22	89.83	2.59	36.02	84.49
If age, time and average PA and diet changed	4.65	64.81	100	3.06	42.63	100
Contribution of change in PA in on weight change	0.48	6.60	10.19	0.47	6.61	15.50
Contribution of changes in Diet in on weight change	0.24	3.31	5.11	0.25	3.49	8.18

Notes: Derived by using observed average covariate values for at the baseline wave and their estimated coefficients, and putting them through Eq (2) from 1991 to 2006. For covariates that changed, these were based off average values.

* Actual weight change between 1991 and 2006 was 7.18 kg

TABLE 6—CONTRIBUTION OF PHYSICAL ACTIVITY AND ENERGY FROM DIETARY FAT ON WEIGHT GAIN AMONG CHINESE MEN

	Observed 15 year change (1991-2006)			**One-time change**		
	(1) Change (1991-2006)	**(2) Coefficient from Table 4, Column 3 or 4**	**(3) Contribution (1)x(2)**	**(4) Average change per year**	**(5) Short run (4)X(2)**	**(6) Long run (4)X(2) / 1-α**
Work & Domestic Physical activity level (METs-hours/week)	-142.75	-0.00307	**0.438 kg (6.1 %)**	-9.5	0.0292 kg	**0.0330 kg (6.9%)**
Diet: Macronutrients			**0.2072 kg (2.9 %)**			**0.0151 kg (3.2%)**
Energy from Fat (%)	6.07	0.174	1.0562 kg (14.8 %)	0.4	0.0696 kg	0.0786 kg (16.4%)
Energy from Carbohydrates (%)	-6.79	0.125	-0.849 kg (-11.8 %)	-0.45	-0.0563 kg	-0.0635 kg (-13.2%)
Diet: Food sources			**0.2727 kg (3.8 %)**			**0.0203 kg (4.2 %)**
Energy from Animal Source Foods (%)	3.60	0.063	0.2268 kg (3.16 %)	0.24	0.01512 kg	0.0171 kg (3.56 %)
Energy from Edible Oils & Fats (%)	6.34	0.089	0.564 kg (7.86 %)	0.42	0.03738 kg	0.04224 kg (8.8 %)
Energy from Fruits and Vegetables (%)	-0.23	0.0178	-0.0041 kg (-0.006 %)	-0.02	-0.000356 kg	-0.0004 kg (-0.08 %)
Energy from Whole grains (%)	-3.15	0.0429	-0.135 kg (-1.88 %)	-0.21	-0.009009 kg	-0.0102 kg (2.12 %)
Energy from Rice (%)	-9.48	0.040	-0.379 kg (-5.28 %)	-0.63	-0.0252 kg	-0.00285 kg (-5.94 %)
Weight	**7.18**			**0.48**		

Notes: Where α is the coefficient of lagged weight (0.114 to 0.118)

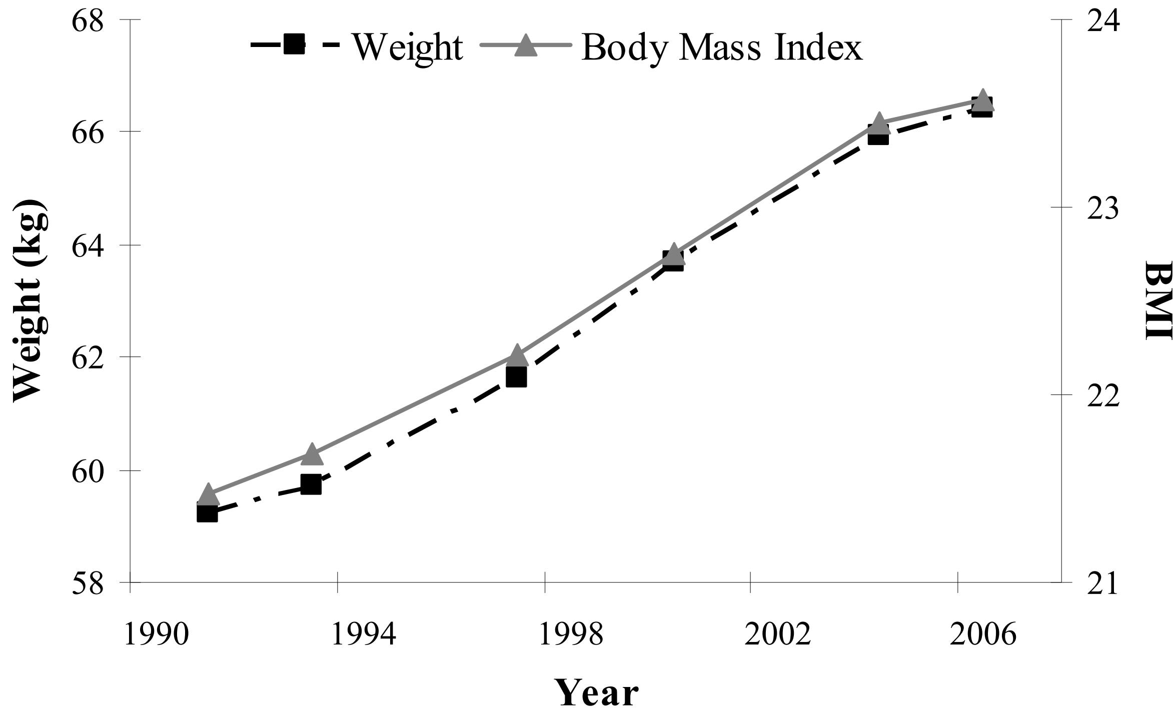

FIGURE 1. WEIGHT AND BODY MASS INDEX OF ADULT MEN IN CHINA FROM 1991 TO 2006

Source: China Health and Nutrition Surveys 1991, 1993, 1997, 2000, 2004 and 2006.

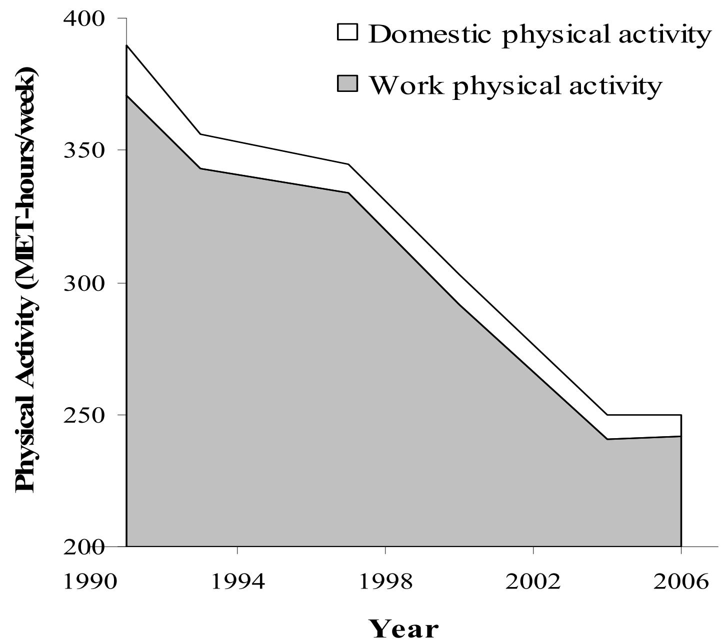

FIGURE 2A. CHANGE IN WORK AND DOMESTIC ACTIVITY

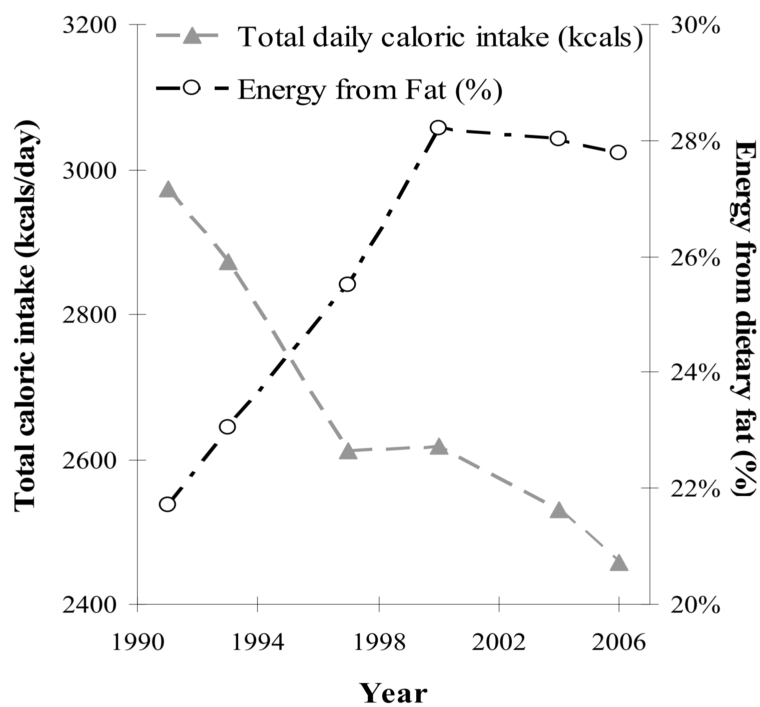

FIGURE 2B. CHANGE IN TOTAL CALORIC INTAKE AND ENERGY FROM FAT

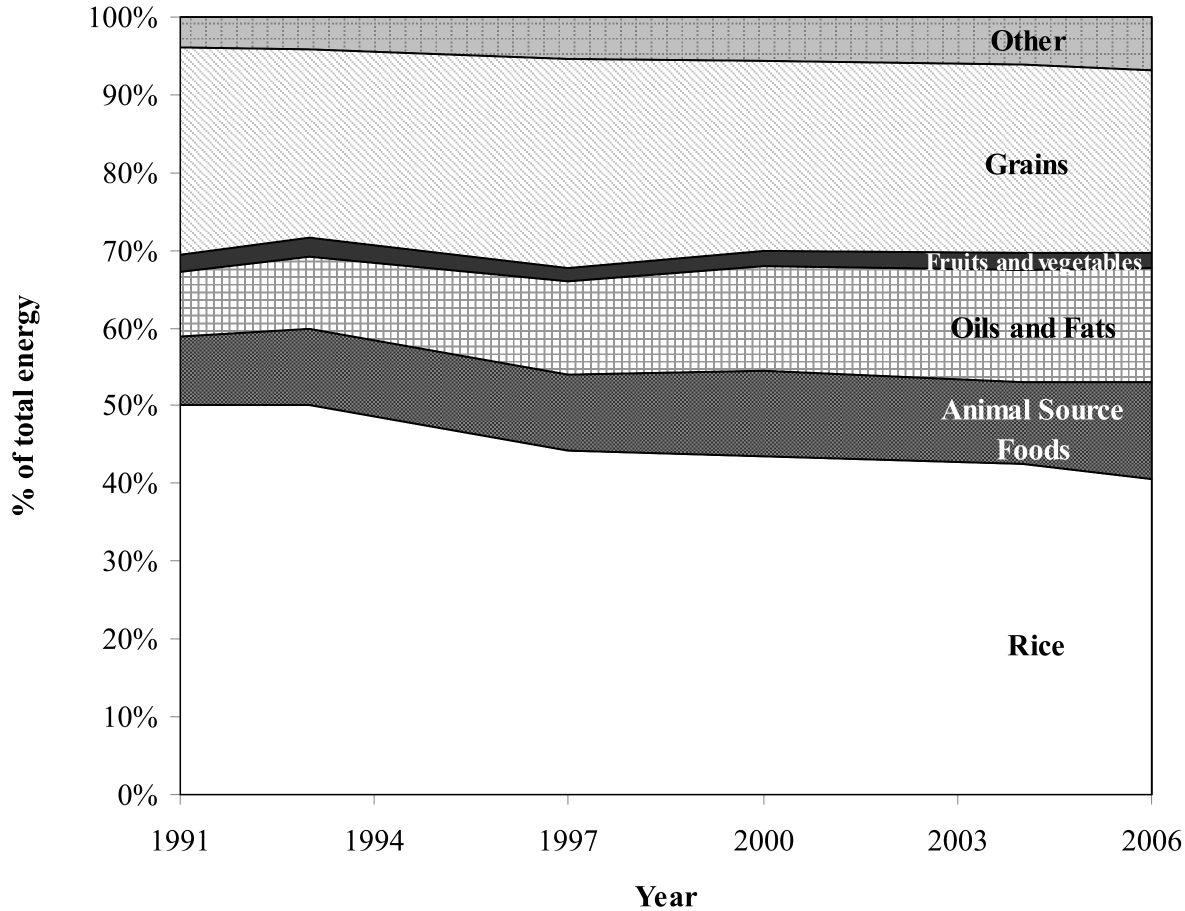

FIGURE 2C. CHANGE ENERGY FROM VARIOUS FOOD SOURCES

Notes: Only among adult (18-55 year old) men.
Source: China Health and Nutrition Surveys 1991, 1993, 1997, 2000, 2004 and 2006.

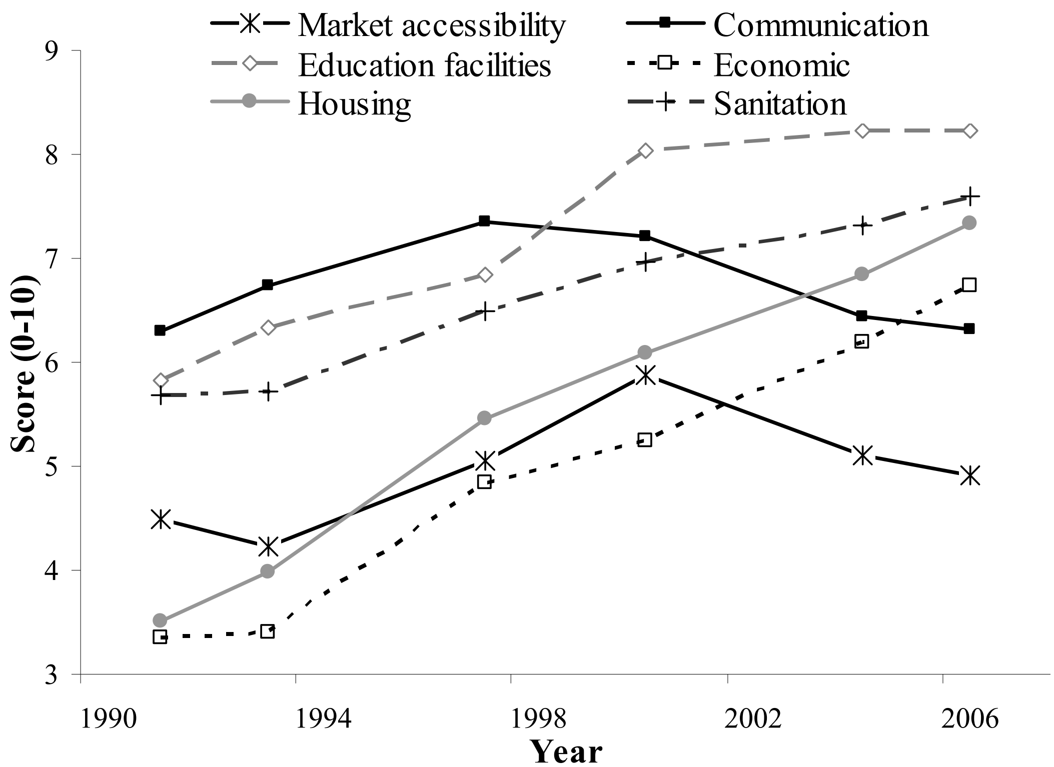

FIGURE 3. AVERAGE CHANGE IN SELECT URBANICITY DOMAINS AMONG COMMUNITIES IN CHINA FROM 1991 TO 2006

Source: China Health and Nutrition Surveys 1991, 1993, 1997, 2000, 2004 and 2006.

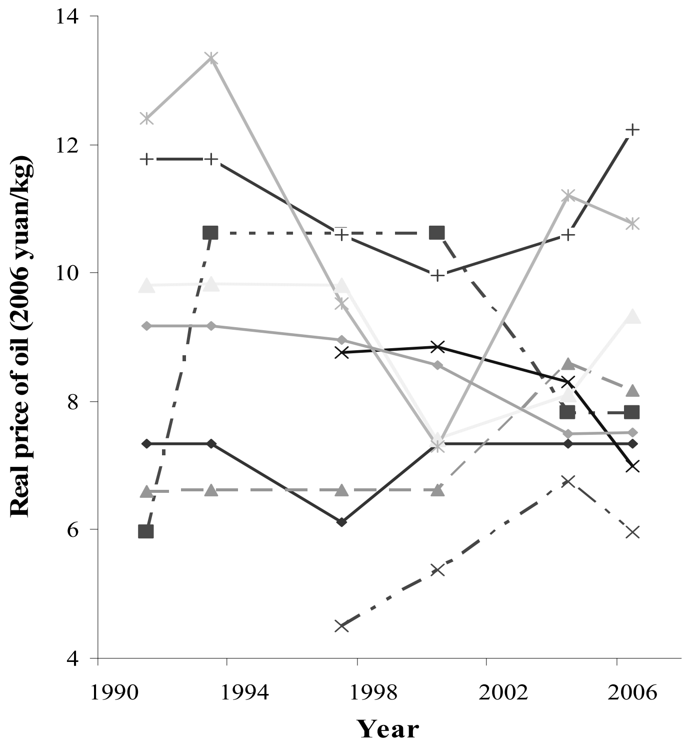

FIGURE 4A. CHANGE IN REAL PRICE OF OIL ACROSS THE 10 MOST POPULOUS COMMUNITIES

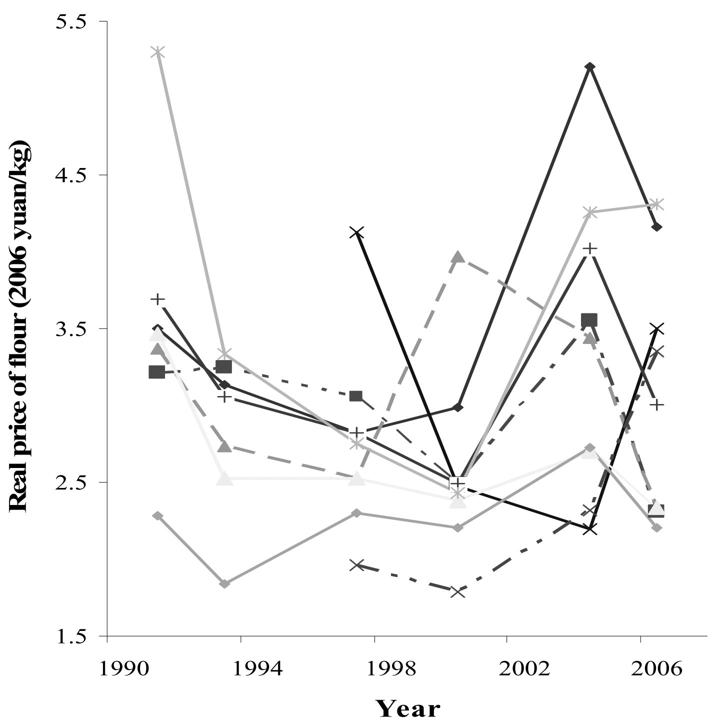

FIGURE 4B. CHANGE IN REAL PRICE OF FLOUR ACROSS THE 10 MOST POPULOUS COMMUNITIES

Source: China Health and Nutrition Surveys 1991, 1993, 1997, 2000, 2004 and 2006.