LONG TERM EFFECTS OF EXPERIENCE DURING YOUTH: EVIDENCE FROM CONSUMPTIONS IN CHINA

By

K. Sudhir and Ishani Tewari

October 2015
Revised March 2016

COWLES FOUNDATION DISCUSSION PAPER NO. 2025R

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Long Term Effects of Experience during Youth: Evidence from Consumption in China*

`

K. Sudhir
Yale School of Management
k.sudhir@yale.edu

Ishani Tewari
Yale School of Management
ishani.tewari@yale.edu

March 2016

* We thank participants at the 2015 TPM Conference in Atlanta, the 2015 Marketing Science Conference in Baltimore, the 2015 China India Insights Conference in New York and the Yale Marketing Group Lunch Seminar, JP Dube and Bart Bronnenberg for feedback. Qiaowei Shen generously provided the fast food outlet data used in the paper. We are grateful for a research grant from Andrew Redleaf of Whitebox Advisors through the International Center for Finance at the Yale School of Management.

Long Term Effects of Experience during Youth:
Evidence from Consumption in China

Abstract

We test for the long-term effects of experience during youth on consumption in non-traditional taste-forming categories. A unique dataset that tracks individuals over twenty years from 1992-2011, residing in nine Chinese provinces that vary widely in both income levels and rate of economic growth, helps us identify cohort and intra-cohort "prosperity-in-youth" (PIY) effects on consumption. We first demonstrate that non-traditional category consumption increases strongly among cohorts that entered adulthood during China's boom years. We then show evidence of the intra-cohort PIY effect, controlling for individual level experience by leveraging the heterogeneity in the timing and rate of growth in prosperity across Chinese provinces. We find that the PIY effect has two dimensions-- a direct effect of one's own prosperity and an indirect effect of the prosperity of one's province during youth. The indirect effects suggest that norms and aspirations created by the consumption of non-traditional categories by the surrounding rich during one's youth have significant impact on long-term consumption—almost the same magnitude as the direct effect. We conduct a large number of robustness checks; in particular, we rule out potential supply side and attitude based explanations for the PIY effect. Our results imply that segmentation and consumption forecasts based on birth cohorts and experience of prosperity can be effective for taste forming non-traditional products in emerging markets.

Key Words: cohort effects, lifecycle effects, emerging markets, China, prosperity in youth, impressionable years hypothesis, long-term effects

"Even if the rest of one's life consisted of one long process of negation and destruction of the natural world view acquired in youth, the determining influence of these early impressions would still be predominant."

Karl Mannheim, "The Problem of Generation" 1923

1 Introduction

Marketers often tout the large numbers of new middle class consumers in emerging markets in describing their market potential in making the case for entry into these markets. Indeed, the estimated size of the middle class for emerging markets such as China (300 million) and India (150 million) dwarf the populations of many large developed economies. In a 2009 article, title "Food Fight," the Economist captured the excitement around selling to the "emerging market billions" well: "Across the developing world, millions perhaps billions of people are currently forming tastes that will endure for the rest of their lives. Put one of Kraft's Oreos or Cadbury's Flakes in their hands and they may become loyal customers for decades to come."

However, for many types of taste based non-traditional products, sales never reach the exuberant expectations at the time of entry into these markets. As an example, despite entering India as early as 1995, with the idea that it would be easy to gain share against traditional Indian breakfast items such as idlis and vadas, Kellogg's had insignificant sales for breakfast cereal in 2011.[2] Similarly, P&G found that despite rising affluence it was not easy to successfully market sanitary napkins to middle class women in India.[3] The reality is that conflict with traditional tastes, values, norms and habits limits consumption for many non-traditional categories, even if consumers can afford them. Marketers of non-traditional

[2] Bolton (2012) in an HBR article notes: "Kellogg's invested $65 million in 1995, establishing an operational and marketing presence to launch Corn Flakes, Wheat Flakes, and its "innovation" — Basmati Rice Flakes — throughout the country with the goal of building a $3 billion business, confident as the managing director of Kellogg India noted: "Our only rivals are traditional Indian foods like idlis and vadas." The article goes on to ask, "How is it possible that Kellogg could envision building a $3 billion business in India, invest $65 million in the first year alone, and end up, 16 years later, with only $70 million in annual revenues? And how can other business leaders avoid making similar mistakes?"

[3] A young Indian woman quoted in a Business Insider article (Srivastava 2014) states: "Although I belonged to a very well-to-do family, we had to use discarded cloth during periods, which we had to wash and reuse. It was not about affordability. It was because of the shame associated with buying sanitary napkins."

categories therefore need to identify consumer segments who can both afford the product, and are also willing to adopt new attitudes and habits. One of the important shifters of tastes, traditions and norms over time are generational cohorts. Cohorts are groups of individuals who share experiences due to their *common* age at any point in time. Defining events such as economic changes, technological progress, socio-political climate, or conflict impact people of different ages differently, but create common attitudes among those in the age group. A particularly interesting age group is youth. In social psychology, the "impressionable years hypothesis" (Krosnick and Alwin 1989) states that individuals are most open to forming new attitudes and change in attitude during late adolescence and early adulthood and common experiences create enduring attitudes and preferences.[4] Thus, cohorts exposed to similar conditions during these formative years develop common attitudes and values. This can lead to similar consumption decisions-- especially in taste based products like music and food.

The fast pace of change in emerging markets such as China or India implies that even adjacent cohorts will experience dramatic differences in their economic, social and political environments during youth. This can lead to significant cohort effects among people who reached adolescence and adulthood as these countries embraced market reforms and experience rapid economic growth. Even within a cohort, there is immense variation in social and conomic conditions depending on location in different provinces in these large countries. This implies, that there is likely to be substantial heterogeneity in experiences during youth even within cohorts, leading to important intra-cohort effects. Growing incomes and variation

[4] As Giuliano and Splimbergo (2014) state summarizing several studies in social psychology, the historical environment during the impressionable years shapes the "basic values, attitudes, and world views of individuals." An anatomical explanation is provided in Glenn (1980) who describes the adolescent brain as in a transitional period with different anatomical and neurochemical features than the adult brain. In particular in this transitional state, "*the volume of grey matter in the cortex gradually increases until about the age of adolescence, then sharply declines as the brain prunes away neuronal connections that are deemed superfluous to the adult needs of the individual.*" In a similar vein, the "increasing persistence hypothesis" proposes that people become gradually more resistant to change throughout their lives. They experience a decrease in flexibility and responsiveness to the wider social environment around them due to a "decline in energy and loss of brain tissue, to disengagement and a decrease in interest in events distant from one's immediate life, and to the accumulation of friends who share similar world views."

in these incomes across provinces is the hallmark feature of emerging markets. Thus, in studying intra-cohort effects, we will specifically focus on the experience of this prosperity or economic well-being during the impressionable years on consumption. We label this the "prosperity in youth" (hereafter "PIY") and show that it can serve as a basis of segmentation and is a useful predictor of long-term consumption for new categories that take off in response to economic growth.

We postulate two mechanisms through which experiencing prosperity during youth impacts long-term consumption. The first is a direct effect through one's own youth-time prosperity. Higher personal income during youth allows a consumer to buy and consume the non-traditional product when attitudes and habits are being formed. This past experience can impact consumption over the long-term. A second indirect effect is through prosperity in one's surroundings during youth. Prosperity in one's surroundings increases the general consumption of non-traditional categories, which in turn can create favorable norms, attitudes and aspirations towards these products and impact their consumption in the long run. In fact, the indirect effect can occur for a consumer even in the absence of a direct effect; i.e., even if one cannot afford and consume a non-traditional product in one's youth, the favorable attitudes formed due to the surrounding prosperity during one's youth can have a positive impact on future consumption, when the individual becomes becoming rich later in life.[5]

Throughout the paper we emphasize that it is critical to distinguish cohort effects from "life-stage" effects. That is, the propensity to buy new "taste based" categories attributable to the cohort effect does not change as a function of age. The challenge of decomposing age, period/time and a cohort effect has confronted scholars across different fields (often called an APC decomposition, which we describe in more detail in the next section) and is often

[5] In discussing the role of product placement in glamorous Chinese movies such as Tiny Times (a Chinese version of Sex and the City), Bloomberg BusinessWeek quotes a Shanghai college student about the potential long-term benefit of such placement among youth who currently cannot afford such luxuries: "*The movie is somewhat detached from our current lives now, and we can't afford a lot of those luxuries. But some of the brands we may buy as we get older.*"

intractable without strong normalization assumptions. In addition to the cohort effect, we are able to identify the intra-cohort PIY effect by exploiting another level of variation with cohorts—economic variation across geographies. However, estimating the intra-cohort PIY effect comes with heavy data requirement, even beyond what is needed for a basic APC –type analysis. First, this requires rare-to-find individual level panel data on consumption for a large number of years that span multiple life stages of an individual. Second, testing for the direct PIY effect requires that there has to be substantial variation in the level of prosperity experienced across and within individuals over time. Third, as individuals within a birth cohort share a large number of experiences—economic conditions, technological progress, cultural norms and many other unobservable characteristics, it is not normally possible to identify the indirect PIY effect at the level of each province, unless there is variation in the timing of when different provinces experienced prosperity.

China provides an ideal setting to empirically test for the PIY effect. As a rapidly growing emerging market, it not only has a high overall rate of income growth, but there is also substantial heterogeneity in the levels and rate of income growth across individuals and provinces. Figure 1a shows the GDP per capita for the different Chinese provinces in the sample. It is clear that not only are the levels different, but the growth rates are substantially different across the provinces as well. This makes it feasible to test for the direct and indirect PIY effect in China. Second, we were able to access rich micro-data-- a very long panel dataset of individual consumption for Chinese individuals.

The China Health and Nutrition Survey tracked the consumption, preferences and health outcomes of over 8000 households from nine Chinese provinces over a twenty-year period from 1992-2011 allowing us to study the panel consumption behavior of over 100,000 individuals. The survey collected detailed information on not only household consumption across multiple non-traditional and traditional categories, but also the changing income levels and other demographics of these individuals. Thus we are able to control for a rich set of contemporaneous factors, demographics, life stage (age) and cohort effects in isolating the PIY effects in a robust manner. The data on consumption and preferences for non-traditional categories over time in which we should not expect PIY effects, provides us an opportunity to conduct falsification tests. Finally, we obtain additional data on the number of fast food outlets across different provinces which is one of the tests that allows us to rule out plausible supply side explanations in isolating the PIY effect.

Our key findings are as follows. First, we find strong evidence of a cohort effect--"millennials," born during the nationally high growth decades (1980s and 1990s) are more likely to consume non-traditional categories like coffee, western snacks, and fast food. In fact, once we cohort effects are taken into account, age no longer becomes an important predictor of purchase! Second, we find evidence of the "prosperity-in-youth" (PIY) effect; individuals who experience direct prosperity in youth or spent their youth in richer provinces are more likely to consume new categories in the future—even 20 years later. Importantly, the intra-cohort PIY effect absorbs all of the cohort effects, making it an even more powerful and effective segmentation construct than pure cohort effects, in that it can leverage geography, income and time more effectively for segmentation. Our estimates imply that if a person spent their youth in coastal Jiangsu (the richest province in our sample) instead of in Western Guizhou (the poorest), their likelihood of coffee consumption would be more than doubled. Third, disentangling the direct and indirect channels of the PIY effect, we find that although higher incomes at both the personal (own-income) and societal (provincial income) level are significant predictors of consumption in non-traditional categories, the effect of the latter is larger in magnitude. Further, the indirect PIY effect affects all individuals, but the effect is almost double in magnitude among individuals, who had low incomes during their youth and therefore could not afford these categories during youth. This suggests that indirect effect creates aspirations for future consumption among the poor to consume non-traditional categories when they became richer. Finding this disproportionately larger effect on poorer individuals also helps us rule out supply-side explanations. That is, the PIY effect cannot be driven solely by availability if poorer people or places, typically also associated with low supply, are the ones who are impacted. Our argument against plausible supply side effects is further strengthened by using additional controls for number of outlets.

Falsification tests build confidence in the PIY results. First, the PIY effect does not hold for traditional categories. Second, a falsification (placebo) test that assigns individuals "fake" ages yields no significant correlation between past prosperity and adoption, suggesting that the prosperity during youth is truly the underlying explanatory factor driving the results.

These results have broader implications for marketing in emerging markets. First, the fast growth of the past two decades in China will have persistent effects on current non-traditional category consumption. Second, given the considerable heterogeneity in growth across different provinces of China, the propensity to consume non-traditional categories will differ for people of the same age if they live in different locations during their youth. This suggests that the market potential will differ by consumer age and geography and should be used for effective segmentation and targeting. Further, the finding that reaching the impressionable youth demographic can have long-term benefits for sales via the direct and indirect channel can guide marketing programs. For youth consumers whose new category adoption over the long-run is driven mainly by affordability, low price points will allow them to try out new and expensive categories during this taste forming and malleable period. On the other hand, affordability is not the sole driver of purchase-- the indirect effect or "aspirational" effect on poor consumers through the aspirational channel suggests that marketers may want to invest in aspirational advertising for these non-traditional categories in richer provinces to shape long-term attitudes and future consumption.

The rest of the paper is organized as follows: Section 2 discusses the relevant background literature. Section 3 describes the data and Section 4 describes the analysis. Section 5 concludes.

2 Related Literature

Our analysis of the cohort and the intra-cohort PIY effect relates to and contributes to four areas of existing research. First, we augment a small, recent literature that explores the effect of an individual's past experience on current consumption. Bronnenberg et al. (2012) and Atkin (2012) use data on consumption of migrants in US and India respectively to show that the purchases/consumption of migrants continue to be impacted by the consumption patterns in states from which they originated—even several years after they migrated to new states. Kuneng and Yaklolev (2014) demonstrate cohort effects in vodka/beer consumption tied to habit persistence and product availability. Beer became available in Russia more freely after 1989 (breakup of the Soviet Union); cohorts that came of age before 1989 and had formed drinking habits by then continue to consume vodka at higher rates, while those who came of age after 1989 consume beer at higher rates. Eizenberg and Salvo (2014) find that Brazilian households exhibit persistence in the soda brands they consume (premium or private label) based on what they chose when they first entered the middle class and started drinking soda. Older cohorts faced large price differentials between national and private labels when they first entered the middle class. They therefore started with cheaper private labels. They continued to buy the private labels at higher rates than the younger consumers who entered the middle classes after the price differentials were reduced.

Each of the above studies focus on one dimension of past experience (i) geography (Bronnenberg et al. 2012 and Atkin 2012), (ii) time (Kuneng and Yaklolev 2014) and (iii) income shift (Eizenberg and Salvo 2014). The PIY effect encompasses all three of the dimensions, and shows that "where" (geography) and "when" (a particular point of time in one's life) one attains adulthood and prosperity can have long-term effects on consumption. Therefore, income and geography at a point of time in one's life (youth) are a valuable joint basis for segmentation in high-growth emerging markets. Its practical value as a segmentation basis is particularly high in emerging markets, where markets and consumer characteristics are changing rapidly and heterogeneously across geographies.

Second, our work is related to a broader literature in social psychology, political science and economics about how susceptibility to attitude changes over an individual's lifecycle. There are two key hypotheses about susceptibility to new attitudes over a life time; (1) the life stages hypothesis, and (2) the impressionable years hypothesis. The life stages hypothesis suggests that openness to different attitudes vary over age; for instance youth tend to be more liberal and become more conservative as they become middle aged. This is well summarized by the saying: "If you are not a radical at twenty, you have no heart; if you are still one at forty, you have no head." The consumption related implication is that that differing needs drive consumer needs at different life stages. For example, financial products tend to have needs at different life stages (e.g., Li et al. (2005). One has very simple financial needs at eighteen (checking account/credit card), but much more complex financial needs during middle age (mortgage, insurance etc.), and again simpler needs when older (e.g., retirement products). The "impressionable years" hypothesis, in contrast and will be our preferred hypothesis in this paper due to the nature of the categories we study, suggests that an individual is most open to attitude change during youth, and subsequently retains these stable attitudes over their lifetime. In terms of consumption, these are likely true for many taste-based products such as music, movies and food. Our estimated age and cohort effects can shed insights on which of these two hypotheses are supported by the data.

Empirically, several papers have found support for the impressionable years hypothesis in multiple settings, though not in the domain of consumption. Schuman and Corning (2012) find evidence for the "critical years" (essentially equivalent to the impressionable years) phenomenon that national and world events (e.g., WWII, the Vietnam War, fall of the Berlin Wall) experienced during later childhood, adolescence, and early adulthood wield disproportionately larger effects on memories, attitudes, and actions in later life. It has been also used to explain political attitudes, voting behavior, political socialization, and partisan realignment in political science (e.g., Krosnick and Alwin 1989; Sears and Funk 1999; Osborne, Sears and Valentino 2011). Giuliano and Spilimbergo (2014) exploit cross-regional variation in occurrence of recessions during youth to show evidence that this makes individuals more favorable towards redistributive policies; they also provide an extensive review of the literature relating the effect of macroeconomic conditions during particular life-stages on economic outcomes.

Third, our paper is related to the literature on identification of cohort effects. Membership in a cohort implies sharing experiences such as macroeconomic episodes during the same period of their lives or viewed from another angle, going through common life-cycle events such as marriage at similar points in time. In trying to ascertain whether the economic environment specific to a cohort creates preferences that persist over time, this paper is in addition estimating an intra-cohort effect. The well-known econometric challenge is that cohort, life-stage (age) and time are linearly correlated (Heckman and Robb 1985). The basic problem arises because there is a linear relationship between these variables: age equals period minus cohort (year of birth). A common approach to break this linear dependency is to make a normalization assumption by bunching people of multiple birth years into one cohort on the grounds that they share common experiences, which are the basis of the cohort effect. In this paper, similarly we assume that households born in the same half-decade have the same cohort effects.

The empirical marketing work relating to cohort effects is limited. The seminal papers in this area are Rentz, Reynolds and Stout (1983) and Rentz and Reynolds (1981, 1991), who study the impact of changing US age distribution on soda and coffee consumption respectively using repeated cross-sectional data. They circumvent the linear dependency identification issues using the normalization assumption of combining multiple age classes into one cohort. The key conclusion in Reynolds and Stout is that contrary to a simple reading of the age profile for soda consumption, which might suggest that soda consumption would decline as the US population aged, the cohort analysis showed that soda consumption would not decline (an insight that proved to be true). Rentz and Reynolds (1991) use the cohort analysis to forecast coffee consumption more accurately than a cross-sectional analysis. Fukuda (2010) applies a Bayesian cohort model to study age, period and cohort effects for expenditures on vehicles. In this paper, we go beyond cohort identification to also identify an intra-cohort PIY effect. This is feasible given that we have detailed individual panel data over a very long period of 20 years, and there is substantial intra-cohort variation in China due to the cross-regional variation in differential income growth across Chinese provinces and own income exposure at youth at the individual level.

Finally, our paper is related to the literature on new product adoption that explores the relationship between personal characteristics and new-product adoption behavior. This literature finds that "innovative" consumers can be characterized by socio-demographics, especially age, income and education (Gatignon and Robertson 1991, Dickerson and Gentry 1983, Im and Bayus 2003). Note that we control for current characteristics, but in contrast to these streams of work, our emphasis is on past experience during youth rather than current characteristics of individuals in influencing adoption. We postulate that past experiences may in fact be equally or more important relative to the influence of contemporaneous observables on adoption/consumption. Further, our work suggests that "innovativeness" as a trait may be explained by past experience of prosperity and geography at a particular age.

3 Data and Descriptive Evidence

We begin with a description of the data and then provide some descriptive patterns in the data.

3.1 Data

We use the China Health and Nutrition Survey as the primary source of our data as ana analysis. It is a unique panel dataset that tracks individuals over two decades over nine waves of surveys. Specifically, we have data from 1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009 and

2011.[6] A major advantage of this dataset is that it spans China during a period of enormous spatial and temporal change. The data cover nine provinces that vary substantially in geography, current as well as historic economic development, and public resources. The map in Figure 1(b) shows the CHNS provinces on which we base our analysis shaded in dark green. The provinces along the east coast (Jiangsu, Liaoning and Shandong) are more advanced and have higher GDP levels and growth rates relative to the inland/northeastern provinces (Guanxi, Guizhou, Heilongjiang, Henan, Hubei and Hunan). We use all the waves except the first one (as information on category consumption was not collected in the first wave). We keep individuals aged 15-85 in every wave giving us an unbalanced panel of 30903 individuals across 8394 households.

We report the descriptive statistics in Table 1. The key dependent variable for our analysis is consumption of three relatively "new" non-traditional categories in China: coffee, fast foods (like KFC and McDonalds) and salty snack foods (like potato chips, pretzels, French fries). Consumption is a dummy variable indicating whether an individual drinks or eats that category during the time of the data collection. Additionally, for coffee, we also construct a share which measures how much coffee s/he drinks relative to tea (i.e. amount of coffee divided by sum of coffee and tea). For our main regressions, we use all three categories. However the length of our panel is shorter for fast food and snacks; so we focus on for which we have the longest panel of data, when we need to further sub-segment markets (as we do when we test for the aspirational mechanism). We also consider measures of preferences as a dependent variable measuring consumption to assess the robustness of the results. The survey asks individuals how much they like fast food or snacks. If they dislike them very much or somewhat we code this as "low preference" and if they like them very much or somewhat then we code this as "high preference."

[6] The survey is a collaborative effort between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. Details are at http://www.cpc.unc.edu/projects/china.

We construct our main independent variables i.e. the measures of “prosperity”—(i) provincial income during ages 18-28 and (ii) personal income during ages 18-28 as follows. For the former, we use macroeconomic data from the EMIS Emerging Markets database. This database provides time-series on the gross domestic (province) product per capita across China. Given our long panel, we know where an individual lived when they were 18-28. We construct an average of the provincial income across all available years for an individual in this range. We construct the personal household income during ages 18-28 from the information in the panel itself.[7] Older individuals who were older than 28 when the survey started will not have an observation for personal income. Figure 2 shows how many people we observe in each wave and what the maximum age in every year. For example, in 1993, only respondents 28 or younger will have a value for personal income while by 2011, those initial 28 year olds would have aged and so the maximum age increases (there is still the possibility of attrition i.e. respondents dropping out or not available for the survey which can lower sample size).

In addition, we have information on socio-demographic and household characteristics like gender, relationship to head, education, urban or rural status. We also have information on contemporaneous household income and contemporaneous province income. The contemporaneous income measures are particularly important to help control for the current “affordability” effect that is likely correlated with prosperity experienced during youth.

Given the time, province in which lived, and age of the individual, we are able to estimate time, province and age fixed effects. In addition, we use information from the birth decade (1940s, 1950, 1960s, 1970s, 1980s, or 1990s) of the individual to isolate cohort effects. Together, these fixed effects on time, province, age, and cohort absorb any variation in

[7] Our choice of this particular age bracket is a slightly longer than some other writing about the impressionable years hypothesis (who typically use 18-25, e.g., Giuliano and Spilimbergo (2014)); but given that our survey is conducted every three years and sometimes an individual may not appear in a particular year, we aggregate income over at least three observations (9 years) rather than just two to have a robust income measure. However, the results do remain robust to the alternative 18-25 year definition.

adoption coming from common national-level shocks, unchanging province characteristics, life-cycle effects and cohort-specific characteristics.

3.2 Model-Free Evidence of Cohort and intra-cohort PIY Effects

Figure 1(c) depicts the growth of non-traditional or Western-inspired categories based on our panel data, which spans the period of rapid economic development in China. We view categories as "non-traditional" if they were not widely available or consumed in China prior to the takeoff of growth, roughly around the early 1980s. Often these categories were commonplace in the West and introduced in Chinese markets by Western firms. For example, coffee has only relatively recently experienced rising acceptance among the Chinese, traditionally a tea-consuming society. Pioneered by Nestle in the early 1990s, instant coffee began gaining popularity and over time, this growing taste has fueled double-digit growth and myriad ways of consumption: coffee shops, fresh grounds, pods—rivaling the variety observed in Western markets. Even though tea still retains a substantially higher market share, coffee volume is climbing steadily.

The aggregate growth rates shown in Figure 1(c) mask substantial heterogeneity in category adoption across age, cohort and geography. In trying to isolate the cohort and geography effect in a model-free graphical manner, we hope to build intuition towards uncovering the presence of a "prosperity-in-youth" effect. The left-hand column of Figure 3 shows coffee, fast food and snack adoption across the age distribution, revealing a mostly monotonically declining profile over age. On first glance, this appears consistent with a life cycle hypothesis of consumption. If coffee is more likely to be consumed earlier in life (starting around age 18), the declining profile over the age distribution is simply due to within-individual variation in consumption over their lifecycle. For example, youth may prefer coffee because it is more fashionable or because it is more potent, but may lose preference for coffee as they grow older.

On the other hand, this particular decline for aggregate coffee consumption over the age distribution can arise due to across-consumer variation alone. Specifically, even if different cohorts have flat consumption profiles across age (and thus no within-individual variation), but if younger cohorts were more likely to consume coffee, then we would generate the declining for aggregate coffee consumption. With cross-sectional data, separating these "age" (within-individual) and "cohort" (across-individual) effects would not be feasible. We leverage our panel structure of the data to disentangle the two effects in Figure 4 where the raw age profile is the solid line. To evaluate the importance of life-cycle driven consumption (within-individual) versus cohort-driven (between-individual), we plot share by age after controlling for individual means or fixed effects as the dashed line We do this by subtracting out average share of each individual from her per-period share and then normalizing the average of the first observed share to zero. So essentially we see the average slope of the age profile over all individuals in the sample after controlling for individual fixed effects. If there was considerable within-individual heterogeneity, then this demeaned profile should retain a positive slope. If, on the other hand, the trend in Figure 3 (or the solid line) was driven by cross-individual heterogeneity, then the profile should be mostly flat. Indeed for coffee in the Panel 4(a), the dashed line looks flatter suggesting no systematic differences in the individual's lifecycle but instead that differences across cohorts seem to be behind the aggregate trends. In Figure 4(b), we plot the same comparative graphs for tea (not a particularly modern category) and see that a very different life-cycle profile from coffee.

To more directly observe cohort effects in consumption habits among Chinese household members, we plot coffee drinking probability by cohort (as defined by decade of birth) in each year from 1993 to 2011, as in the right-hand side panel of Figure 3. Modern categories capture a growing share of an individual's beverage intake especially for younger generations born in the 1980s and 1990s. Individuals born before the mid-1960s (or who came of age before the end of the Cultural Revolution) do not increase adoption over time. For example, younger generations show increasing interest in coffee and are much more likely to be coffee drinkers than the older cohorts both cross-sectionally and over time. Also, the increasing total height of the bars over time indicates that there is an increasing propensity in the overall population to drink coffee.

The evidence so far indicates a possible cohort effect, with younger and impressionable cohorts more likely to consume coffee, fast food and soda, driving the expansion of these new categories. Given our interest on the role of prosperity in shaping preferences, the next question is whether we can link this cohort effect to particular market conditions like economic growth prevailing during youth. The Chinese economy has experienced rapid economic growth, averaging over 9% annually during the past thirty years. As shown in Figure 1(a), this high growth rate is geographically heterogeneous with growth rates generally higher in coastal areas relative to the inland regions. The cohorts born in 1980s and 1990s were in their youthful years during the nationally high-growth period. To probe further into a possible correlation between provincial prosperity and youth, we turn to Figure 5.

Here, we plot consumption in 1993 and 2011 for those born after 1980 (blue) and those born before (red) for high (dashed) and low (solid) income provinces. This definition of "high" and "low" is based on being above and below the median sample level in the top panel. In the bottom panels, we compare the richest (Jiangsu) and poorest (Guizhou) province. Regardless of how growth is defined, the difference in consumption between youth and adults in poor and rich provinces is clear. The steepest slope i.e. the biggest jump in consumption over the time period is for youth in rich provinces and the smallest is for the older generations in poor provinces. In fact, in the top panel, consumption is virtually identical for older generation regardless of their location. In contrast, for the youth, the slope is much higher for the youth in richer provinces. Thus these descriptive model-free analysis shows evidence of "prosperity in youth." We next explore the "prosperity in youth" hypothesis more formally through regression analysis, controlling for both time-varying and time-invariant factors.

4 Empirical Analysis

Next, we describe our model specification and our empirical strategy to identify cohort and PIY effects. We then perform robustness analysis and falsification checks. Finally, we explore the mechanism underlying the PIY effect.

4.1 Model Specification

We begin by estimating the following specification to isolate the cohort effect on consumption from age, time, province and contemporaneous individual effects:

$$Category_{ipt} = \beta X_{it} + \varphi_c + \alpha_a + \delta_t + \lambda_p + \varepsilon_{ipt} \tag{1.1}$$

where $Category_{ipt}$ is a dummy variable indicating whether individual i in province p consumes the category at time t. $Category_{ipt}$ is appropriately redefined when the dependent variable is preference for the category; or share of coffee relative to coffee and tea. In (1.1) above, X_{it} is a vector personal characteristics which may be time-varying (age, contemporaneous incomes, educational status) or time-invariant (gender, relationship to head, income during youth, rural/urban). We capture the cohort, age and time fixed by φ_c, α_a and δ_t respectively. The time fixed effect captures common national shocks. We control for unchanging characteristics of a province through province fixed effects λ_p. We estimate the regression using the probit model when the dependent variable is consumption and OLS when the dependent variable is preference or share. The reported standard errors are clustered at the individual level.

We next add the direct and indirect intra-cohort PIY effect by including the individual i's own income during youth (18-28) and the per-capita GDP of the province in which the individual lived during youth (18-28) below:

$$Category_{ipt} = \pi_1 \log(I_i^{18-28}) + \pi_2 \log(\text{GDP-pc}_{ip}^{18-28})\beta X_{it} + \varphi_c + \alpha_a + \delta_t + \lambda_p + \varepsilon_{ipt} \tag{1.2}$$

As discussed earlier, we treat individuals born in the same half-decade as belonging to the same cohort. The PIY effect to be estimated in (1.2) relies on intra-cohort variation and is identified off the cross-province variation in prosperity which impacts differently even individuals within cohorts.

4.2 Time, Age, Cohort Effects

We begin by reporting the results of the estimation equation (1.1) for the three categories. The estimates are presented in Table 2.[8] The estimated age, cohort and time fixed effects are shown graphically in Figure 6 for ease of interpretation. First, the time effects displayed in panel (a) of Figure 6, show that consumption in all three non-traditional categories is rising over time on average across provinces. This is not surprising, given the overall economic, and social attitudes towards non-traditional consumption.

Second, the age or lifecycle effects reported in panel (b), shows a monotonic decline, but the decline is relatively flat across all three categories. In fact, the confidence intervals around these estimates are not statistically significantly different from zero, which is the difference relative to age bracket 15-20. This supports the hypothesis that there are no significant life cycle effects in these taste-based product categories. This is in contrast to the earlier age profile from the raw data that did not include time and cohort fixed effects (see Figure 2), where the decline with age was very sharp. These results highlight that a naïve model of age and consumption when used to forecast and segment emerging markets can be highly misleading.

Finally we discuss the estimated cohort effects, shown in Panel (c). The omitted category is the 1940-45 birth cohorts, so the effects reported are relative to this cohort. In contrast to the age effects, the cohort effects are significantly different from zero at least for later cohorts. Across all categories, there is a strong pattern of increasing coefficients on later

[8] The results presented are for the probit. We tested and replicated the main results for the Linear Probability Model.

birth cohorts. The estimates imply that being born in the early 1990s increases coffee drinking probability by 5% and in the late 1990s by 12%. For categories like snacks or fast food, cohorts beginning in the late sixties and late seventies, who came of age in the eighties and nineties begin to display higher consumption likelihoods, which continues to increase for even younger cohorts.

Finally, in Table 2, we see that contemporaneous income has a strongly positive and significant effect on consumption across all categories. This is not surprising and has face validity in that we expect richer individuals to be more likely to consume non-traditional categories, which are generally priced higher than traditional foods.

The basic message from the estimates in Table 2 and Figure 4 is that younger cohorts are more likely to adopt modern categories, but there are very little lifecycle effects over age. Next, we seek to parse out the cohort effects, by assessing the role of China's increasing prosperity in driving and the experience of it during youth-when tastes are being formed. We therefore explore this link between prosperity in youth and consumption.

4.3 The "Prosperity in Youth" (PIY) Effect

We report the results of our specification (1.2) to explore the PIY effect in Table 3. Table 3 shows the coefficients on of one's personal income ("direct effect") during ages 18 through 28 as well as the average provincial income during ages 18 through 28 ("indirect effect")—on adoption of new categories. Results for each of the three categories are in separate panels (a), (b) and (c) for coffee, fast food and snack adoption respectively. Each of the columns adds additional controls or fixed effects.

Column (1) in all panels shows strong support for the hypothesized direct and indirect PIY effect. We see that (i) the direct effect of one's personal income during the youth years is always positive and statistically significant; and (ii) the indirect effect through provincial income at 18-28 is also positive and statistically significant. Interestingly, the indirect effect is almost double the direct effect in terms of magnitude for all three categories.

Next we look across columns to see how the PIY estimates change with additional controls. First, we control for contemporaneous income in Column (2). Including current income tends to reduce the direct PIY effect in all three categories. This should not be surprising due to the expected correlation between an individual's past and present income. Surprisingly, however, the coefficient on the provincial income remains stable (or even becomes stronger) even with the inclusion of contemporaneous personal or provincial income. Interestingly, the addition of cohort fixed effects changes the magnitude of the PIY effect very little. However it increases the standard errors of the indirect PIY effect. This suggests that the intra-cohort effects captured by PIY captures most of the cohort effects. We validate this in panel (d), where we report the pseudo R^2 of adding cohort effects in addition to PIY. There is little explanatory power beyond PIY in cohort effects. This suggests that cohort effects are almost entirely be explained by the PIY effect. Finally, in column (4), we add the age fixed effects. This has little impact on the direct effect for all categories. For the indirect effect, we find little effect on coffee and fast food, but the magnitude of the indirect effect increases for snacks.

We now interpret the PIY effects based on estimates in column (4), our most complete specification. A 10% increase in per capita GDP in an individual's province while they are young, boosts coffee drinking probability by .001 percentage points and a 10% increase in own income boosts this by .0004 percentage points. The mean coffee drinking probability in the sample is .025, so this implies increases of 4% due to the indirect effect and 2% due to the direct effect. Given that China was growing at around 10% GDP rates every year for almost three decades, the effects of income on cumulative coffee consumption over time can be very significant indeed. Specifically the estimated 4% (2%) increase in coffee consumption per year due to 10% income growth compounded over 10 years translates to a 48% (22%) increase in consumption in 10 years. The province effect for snacks and fast food are also roughly the same magnitude.

As we would expect, contemporaneous measures of prosperity i.e. an individual's current household income and the current province income continue to affect consumption of all three categories. Coefficients on other individual characteristics like gender, education status and rural residence show that being male, a student and residing in urban areas is also positively correlated with adoption of modern categories. The rural dummy is strongly negative and significant, reflecting the widely acknowledged gap in socio-economic conditions between rural and urban areas in China—that also impact consumption.

Next we examine the fraction of variance explained by the different set of variables. The results are reported in panel (d). Column (1) reports the pseudo-R^2 with only demographics and province fixed effects. Columns (2)-(5) of panel d show the pseudo-R^2 corresponding to the models estimated in Columns (1)-(4) in panels (a)-(c) of Table 3. Columns (2) and (3) show that the direct effect and indirect effect of PIY add additional explanatory power. In terms of relative share of the variance explained by the direct and indirect effects on consumption, the indirect effect explains 22%, 48% and 48% of the total PIY effect for coffee, fast food and snacks respectively. Column (4) shows no incremental explanatory power relative to Column (3) demonstrating that the intra-cohort PIY effect has explained virtually all of the cohort effects. This provides empirical evidence that PIY can explain all of the cohort behavior in Chinese consumption of non-traditional goods.

Finally, we provide a measure of the economic magnitudes of the PIY effect by reporting how much of the per-capita consumption of the different categories is explained by the PIY effects. To do so, we report in panel (e) how a 1% change in one's own youth income or provincial per-capita GDP at youth will reduce average consumption. For coffee, a 1% reduction in own income at youth reduces average adoption rates by 12% while a 1% reduction in one's provincial GDP at youth reduces average adoption rates by 18%

respectively.[9] The joint effect of a 1% reduction in income and provincial GDP at youth is a reduction in average adoption by 28%. The corresponding numbers for fast food and snacks are smaller, but similar in relative magnitudes. These large effects suggest that the PIY effects are not merely statistically significant, but are economically and managerially very significant.

4.4 Robustness Checks

We conduct a series of robustness checks. First, we replicate results so far with consumption data with other dependent variables such as preference for the category and share of consumption. Next, we conduct a variety of falsification checks to assess the robustness of the PIY effect.

Other Dependent Variables: Preferences and Shares

Our main analysis was based upon the consumption dummy (behavior) as the dependent variable. For coffee, we have information on the relative share of consumption of coffee relative to tea and coffee. We can therefore use share of coffee as a dependent variable and repeated the main analysis. We report the results in Table 4a). Our results for the direct and indirect PIY effect on the relative consumption of the non-traditional category (coffee) to relatively traditional one (tea) are similar to the results from the main analysis for coffee adoption.

For fast food and snacks, we also had preference information for the category from the survey. Here, the individual answers on a scale of 0-4 whether she does not consume, dislikes very much, dislikes somewhat, likes somewhat, or likes very much the category in question. We classify the score on this scale as high or low preference. We report the results in Table 4(a). Our results for the direct and indirect PIY effect are similar to the main analysis in both categories.

Province level unobservable trends

---

[9] The total effect is less than the sum of the direct and indirect effect as there is correlation in the variables and the nonlinear probit link between the variables and probability of consumption.

A possible confound for the indirect PIY effect is that unobservable province level factors correlated with income at youth may also be correlated with current income and this may explain the individual's adoption today. For example, if people in a province generally tend to be more modern and progressive both in the past and contemporaneously this could influence both income and adoption. Insofar, this is a time invariant characteristic, it will be taken care of by province fixed effects. However if there are trends that we have not accounted for, these could explain the estimated PIY effect. In Table 5, we control for linear province specific trends to absorb evolving unobservables. Both the direct and indirect PIY effects generally remain significant for all three categories. The magnitudes of the effects are similar to Column (4) of Table 3 for coffee and fast food, but there is a drop in the magnitude of the indirect PIY effect for western snacks. Overall the continued significance and relatively large magnitudes of the effect boosts confidence in the robustness of the PIY effects.

Falsification Checks

To further bolster confidence in our findings, we conduct two falsification tests. First, we test whether PIY matters for traditional categories like fruits and vegetables, where a priori, we should not expect to see them as they are already well-accepted as indicated by the fact that fruits and vegetables are adopted by virtually everyone (over 99%). While, this means that it is not possible to do the adoption test of PIY on these categories, we can still perform the test with the preference data. The results are reported in Table 6. As expected, we find no effect of PIY on the fruit and vegetable categories, supporting our premise that PIY effects are likely to be strong in new non-traditional categories.

Next, we run a placebo test to check whether some other unobservable trends may drive the intra-cohort PIY effects we estimate. To do this we randomize the cohort by assigning random "fake" birth years to individuals instead of their true ones. This implies that we are also assigning them false income-at-youth measures and ages. We then run the basic specification (1.2) for coffee with 100 such false assignments to see if the prosperity effect still holds—suggesting some other unobservable factor drives the PIY results. The distribution of the t-statistic for the coefficients from the 100 regressions is plotted in online Figure 7 The mean t-stat for the "fake" regressions is centered around zero with a mean of 0.113. The lack of significance in these falsification runs makes it unlikely that the PIY results are driven by other unobserved factors.

4.5 Exploring the Mechanism behind the Indirect PIY effect

We conclude by exploring two alternative mechanisms underlying the indirect PIY effect due to provincial prosperity. Why should someone growing up in a rich province at youth, be more likely to purchase non-traditional categories in the future? We explore two plausible mechanisms: aspirations and supply/availability.

Aspirations

We explore the idea that a young person who does not have enough income during youth may nevertheless form aspirations for consuming non-traditional categories when they are in richer provinces where larger numbers of richer people are consuming the product s/he cannot consume today. Such individuals may buy in the category when they become richer later in life.[10] If aspirations drive the indirect PIY effect, it should be stronger among poorer youth. To test this, we segment households using a median split based on their level of prosperity in youth and measure the indirect PIY effects for the richer and poorer segments at youth. If the indirect PIY effect is positive and stronger among the poorer segment, we consider that as supporting the aspirational mechanism.

Columns (1) and (2) of Table 7 report the regression results with coffee adoption and coffee share as the dependent variables, with the rich/poor segmentation based on youth income. We find significant indirect effects of provincial per capita GDP on coffee consumption for both the rich and poor. Further, consistent with aspirational motives, the indirect effect on consumption is stronger for individuals who were poor when they were

[10] "Aspirations, not current income." New York Times August 17th 2012.

young as reflected in the positive interaction effect of provincial GDP per capita and low income at youth.

One possibility is that unobserved factors about the poor (and not the aspirational mechanism) during youth persists across time and hence drives the consumption result. To rule out this explanation, we conduct the following falsification test. We segment the market by current income (not youth income) and replicate the analysis above. If aspirations created by experiences during youth drive the result, the interaction effect of provincial GDP per capita and low current income should not be significant. The corresponding results for Columns (1) and (2) are reported in Columns (3) and (4) of Table 7. Indeed, we do not find significance for the indirect effect if the individuals are classified as rich or poor based on current income. Overall, these set of results support aspirations as the mechanism underlying the indirect PIY effect.

The magnitudes of the indirect PIY effect on low-income and high-income households shed additional insight on the indirect effect. Earlier, we had noted that the magnitude of the indirect PIY effect was double in magnitude to that of the direct effect. Now by splitting the effects of the rich and the poor, we find that indirect effects affect them both, but the poor are affected almost twice as much as the rich. Thus the long run effect of PIY for the rich and the poor are about the same—while the direct effect contributes to half of the PIY effect for the rich, the indirect effect drives most of the PIY affect for the poor.

Supply

An alternative mechanism underlying the indirect PIY effect is that provincial prosperity is likely correlated with supply side variables at the local level such as distribution and advertising and can impact availability and exposure to the products. We provide two pieces of evidence that our results are unlikely to have arisen due to a supply mechanism. First, our previous test for the aspirations mechanism also serves as a falsification test for the supply mechanism. We expect a positive correlation between prosperity and supply and therefore supply factors should likely be more favorable to the rich than to the poor. This means that if the supply factors are the underlying mechanism, then the indirect PIY effect should be stronger among the rich than the poor. The fact that we find the effect is stronger among the poor suggests that supply factors are unlikely to be driving the indirect PIY effect.

Second, we obtained a province-level panel dataset that measures the number of fast food outlets entering China from 1988 to 2007. KFC and McDonald, two of the largest and oldest players in this category, are the firms represented. We run (1.2) with fast food as the focal category and include controls for supply through the log number of outlets in the market at youth and also contemporaneously. If the PIY coefficients remain significant even after inclusion of these supply variables, then we can conclude that PIY is not entirely driven by availability. Table 8 shows that inclusion of these variables strengthens the indirect PIY effect and preference for fast foods, again suggesting that PIY is not moderated by supply factors.

5 Conclusion

In this paper, we explore how a consumer's experiences during youth impact their consumption behavior well into the future, specifically the propensity to consume new non-traditional categories. We identify substantial cohort effects for non-traditional categories among millennials who came into adulthood during the period of rapid Chinese economic growth. In addition, we estimate the intra-cohort "prosperity-in-youth" effect exploiting the variation in prosperity across households and across Chinese provinces. We find that the PIY effect has both a direct and indirect effect. The direct effect, due to one's own prosperity at youth, increases propensity to consume non-traditional expensive categories, and this has long-term impact on consumption well into the future. The indirect component, due to the prosperity in the province in which the individual lives also affects consumption, and surprisingly is almost twice in magnitude relative to the direct effect. For the rich, the direct and indirect components of PIY are roughly equal, but for the poor, the PIY effect is entirely due to the indirect component.

Our results highlight the importance of taking into account cohort effects, as a naïve age-consumption relationship may lead to the misleading conclusion that consumption of non-traditional categories declines with age in China. In contrast, we find that young people begin consumption of non-traditional categories during youth as China has experienced rapid of prosperity recently but these consumption attitudes persist over time. This means that conversion of youth have significant long-term effects that simply do not decay over time/ Interestingly, we find that the intra-cohort PIY effect that we propose in this paper absorbs all of the cohort effects; there is little explanatory power for cohort effects, beyond the PIY effect.

Our findings have several implications for firms introducing new categories in emerging markets. First, a homogenous segmentation strategy in high-growth and heterogeneous emerging markets may not be appropriate as the target segment varies across time and space. Leveraging this heterogeneity can serve as an effective segmentation and targeting tool. Second, targeting the appropriate youth segments today will not just expand current sales but also future sales, as the preferences formed during the impressionable years remain sticky. From a branding perspective, this suggest that cultivating tastes towards a brand early in the lifecycle of a new category may create a first-mover advantage that lasts over a long time (Carpenter and Nakomoto, 1989). Third, richer provinces are even more attractive for targeting, because even for poor, young consumers in those provinces who cannot afford the category, favorable tastes that will persist into the future are being formed today.

Taken together our findings have important implications for firms in rapidly changing and heterogeneous emerging markets. Getting the right estimate of initial market potential and the trajectory for growth is critically important for managers entering emerging markets. Our evidence suggests that the youth who have experienced prosperity directly or indirectly (by living in richer provinces) in their "preference formative impressionable years" need to be over-weighted in estimating market potential for "sticky" taste and attitude products. The right potential and growth forecasts help to set the right expectations for success and commit the right level of resources at the time of introduction. and enables the development of a feasible plan for ramping up commitment and resources over time. Otherwise the entry becomes a failure, leading to disappointment, potential downsizing or even withdrawal, leaving many managers careers damaged irrevocably for no real fault of their own.

While our results support the idea that economic prosperity during youth increases persistent lifetime consumption behavior, future research can delve deeper into the mechanisms. For example, is our identified aspirational effect moderated through the presence of social networks that influence category adoption? We hope this paper stimulates additional research on the long-term effects of experiences during youth on a variety of outcomes of interest to marketers in new settings—e.g., consumption across diverse categories, voting choices and attitudes towards new trends such as sustainability, health and wellness.

References

Atkin, David. "Trade, Tastes, and Nutrition in India." American Economic Review 103.5 (2013): 1629-63.

Bolton, Robyn. “Are You Targeting a Phantom Market?” Harvard Business Review 2012.

Bronnenberg, Bart J., Jean-Pierre H. Dubé, and Matthew Gentzkow. "The Evolution of Brand Preferences: Evidence from Consumer Migration." American Economic Review 102.6 (2012): 2472-2508.

Carpenter, Gregory S., and Kent Nakamoto. "Consumer preference formation and pioneering advantage." Journal of Marketing Research (1989): 285-298.Deaton, Angus S., and Christina Paxson. "Saving, growth, and aging in Taiwan." Studies in the Economics of Aging. University of Chicago Press, 1994. 331-362.

Dickerson, Mary Dee, and James W. Gentry. "Characteristics of adopters and non-adopters of home computers." Journal of Consumer Research (1983): 225-235.

Economist, “Food Fight,” November 5th, 2009.

Eizenberg, Alon and Alberto Salvo. “The Rise of Fringe Competitors In the Wake of an Emerging Middle Class: An Empirical Analysis,” American Economic Journal: Applied Economics (forthcoming)

Fukuda, Kosei. "A cohort analysis of household vehicle expenditure in the US and Japan: A possibility of generational marketing." Marketing Letters 21.1 (2010): 53-64.

Gatignon, Hubert, and Thomas S. Robertson. Innovative decision processes.” In Handbook of Consumer Behavior. Eds Thomas S. Robertson and Harold H. Kassajian. Englewood Cliffs, NJ: Prentice Hall.

Glenn, N.D. (1980), “Values, Attitudes and Beliefs”, in O.G. Brim and Kagan, J. (eds) Constancy and Change in Human Development. Cambridge, MA: Harvard University Press.

Giuliano, Paola, and Antonio Spilimbergo. "Growing up in a Recession." The Review of Economic Studies 81.2 (2014): 787-817.

Heckman, James, and Richard Robb. "Using longitudinal data to estimate age, period and cohort effects in earnings equations." Cohort analysis in social research. Springer New York, 1985. 137-150.

Im, Subin, Barry L. Bayus, and Charlotte H. Mason. "An empirical study of innate consumer innovativeness, personal characteristics, and new-product adoption behavior." Journal of the Academy of Marketing Science 31.1 (2003): 61-73.

Krosnick, Jon A., and Duane F. Alwin. "Aging and susceptibility to attitude change." Journal of Personality and Social Psychology 57.3 (1989): 416.

Kueng, Lorenz, and Evgeny Yakovlev. “How Persistent Are Consumption Habits? Micro-Evidence from Russia's Alcohol Market.” No. w20298. National Bureau of Economic Research, 2014.

Li, S., Sun, B., and Wilcox, R. T. (2005). Cross-selling sequentially ordered products: An application to consumer banking services. Journal of Marketing Research, 42(2), 233-239.

Lin, Liza. “Why Luxury Brands Love China’s Sex and the City,” Bloomberg Businessweek, January 29th 2015.

Malmendier, Ulrike, and Stefan Nagel. Depression babies: Do macroeconomic experiences affect risk-taking?. No. w14813. National Bureau of Economic Research, 2009.

Rentz, Joseph O., Fred D. Reynolds, and Roy G. Stout. "Analyzing changing consumption patterns with cohort analysis." Journal of Marketing Research(1983): 12-20.

Rentz, Joseph O., and Fred D. Reynolds. "Forecasting the effects of an aging population on product consumption: An age-period-cohort framework."Journal of Marketing Research (1991): 355-360.

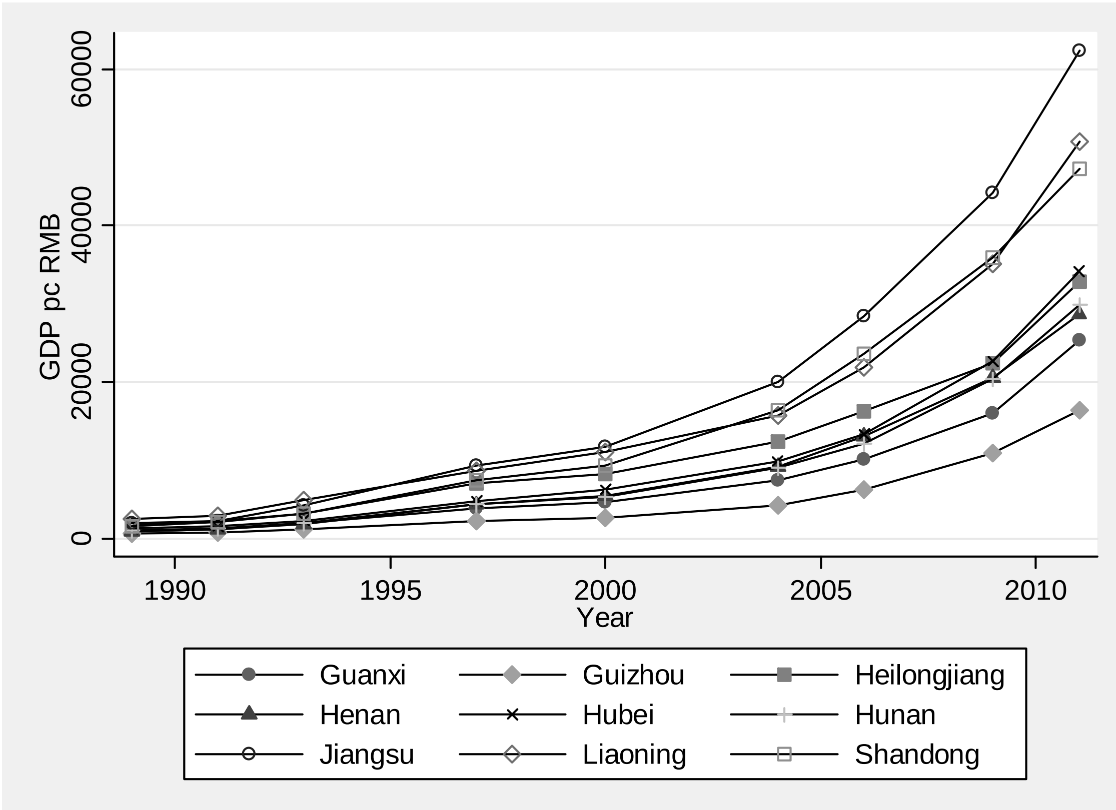
Reynolds, Fred D., and Joseph O. Rentz. "Cohort analysis: an aid to strategic planning." The Journal of Marketing (1981): 62-70.

Srivastava, Abhaya. “Procter & Gamble Wants To Put An End To India's Menstruation Taboos,” Business Insider October 3rd 2014. Schuman and Corning (2012)

Stephens, N. M., Markus, H. R., & Townsend, S. S. M. (2007). Choice as an act of meaning: The case of social class. Journal of Personality and Social Psychology, 93(5), 814–830.

Figure 1: Background on growth and geographic coverage

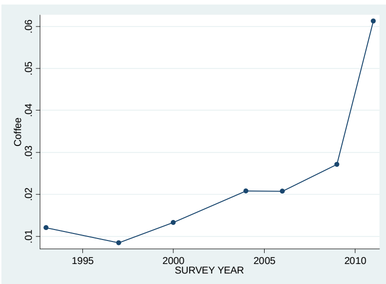
(a) Province Growth 1978-2011



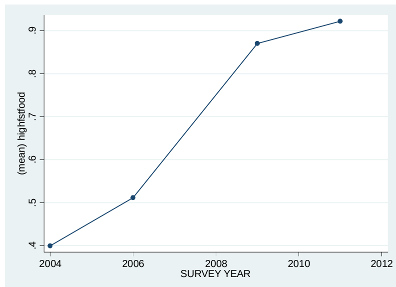
(b) Provinces Covered in CHNS Data (dark green)

(c) Growth of Non-Traditional Categories

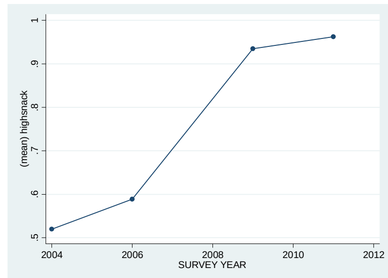
(a) Coffee



(b) Fast Food



(c) Snacks



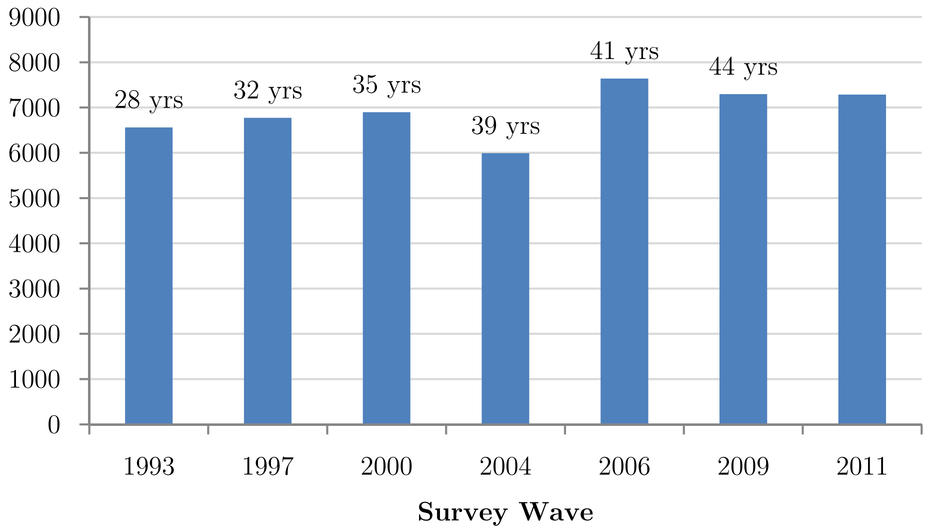
Notes:

(a) depicts the GDP per capita in different provinces of China during the period 1978-2011, with darker shades implying higher growth rates. Data is from the China Statistical Yearbook.

(b) shows the geographic coverage of our data, the dark shaded provinces are the ones included in the sample.

(c) plots the average share of individuals within our sample drinking coffee, and consuming fast food, snacks in the survey year.

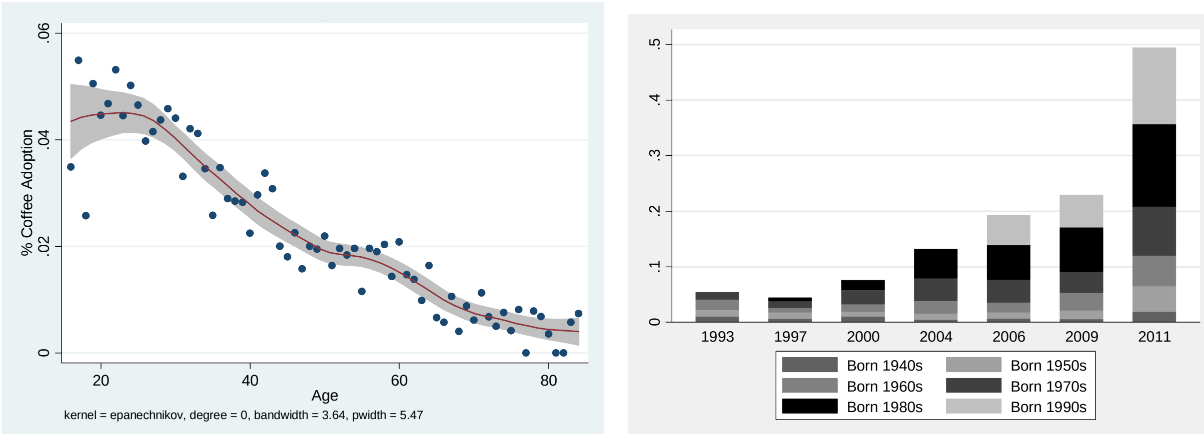
Figure 2: Sample Composition-- Number of Observations and Maximum Age by Survey Year



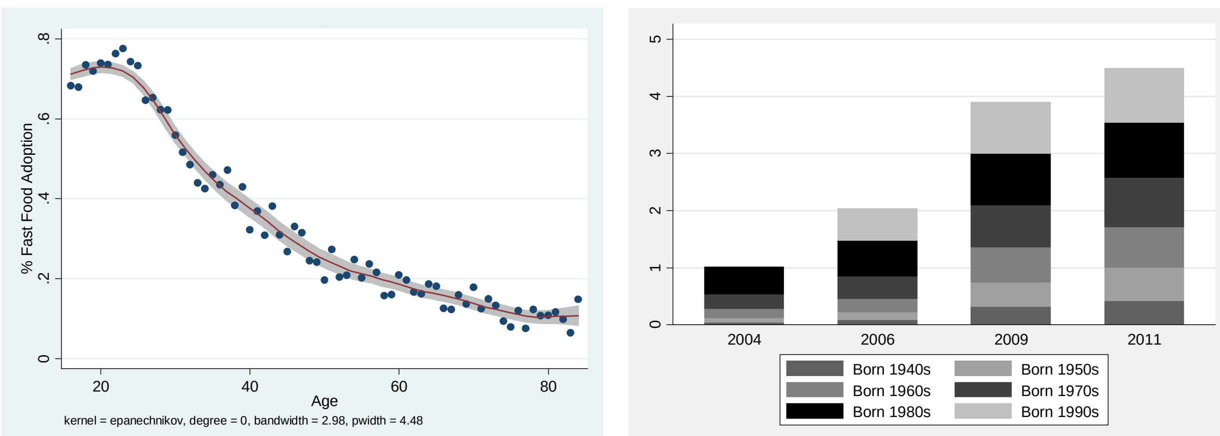
Note: The above figure shows the number of observations per year and the age of the oldest individual in that year i.e. the age below which we have data on youth (18-28) income.

Figure 3: Age and Cohort –wise Category Consumption

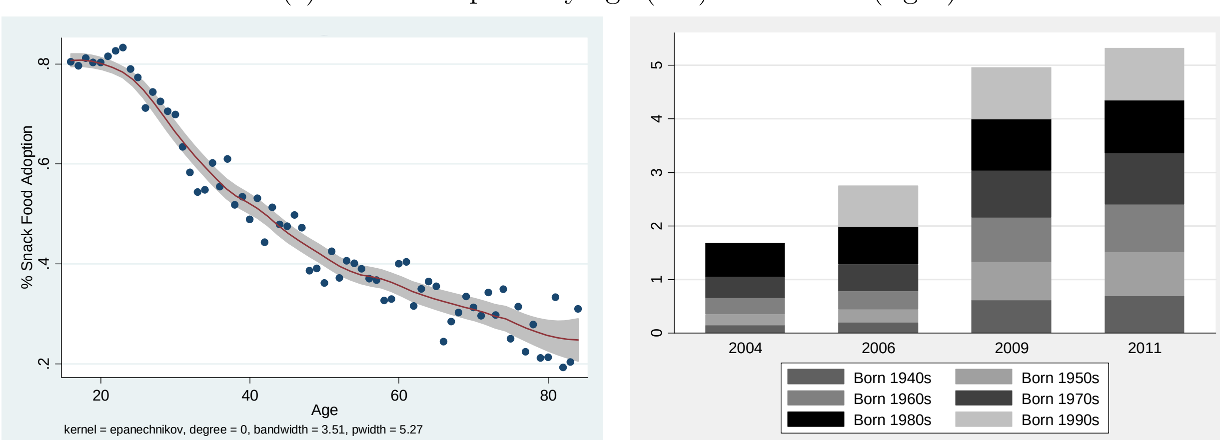
(a) Coffee Adoption by age (left) and cohort (right)



(b) Fast Food Adoption by age (left) and cohort (right)



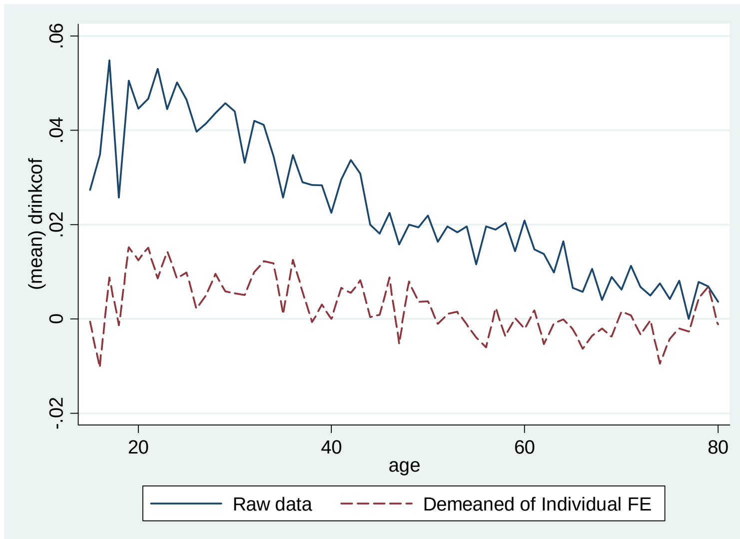
(c) Snack Adoption by age (left) and cohort (right)



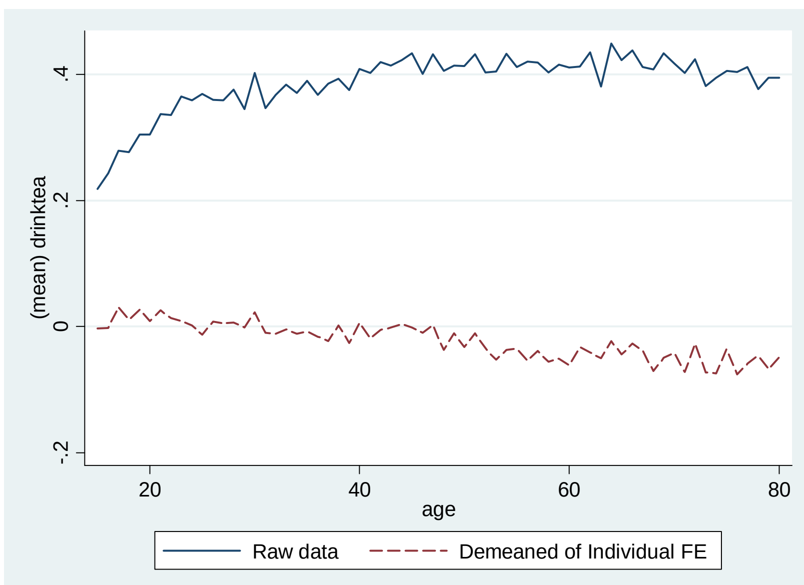
Notes: The left panel figures show the local polynomial-smoothed life-cycle profiles of category adoption by age. The darker shaded bands around the smoothed lowess plot is the 95% CI. The right panels show average adoption within sample by cohort, as defined by decade of birth.

Figure 4: Detangling Across-Individual and Within-Individual Variation Adoption by Age

(a) Coffee



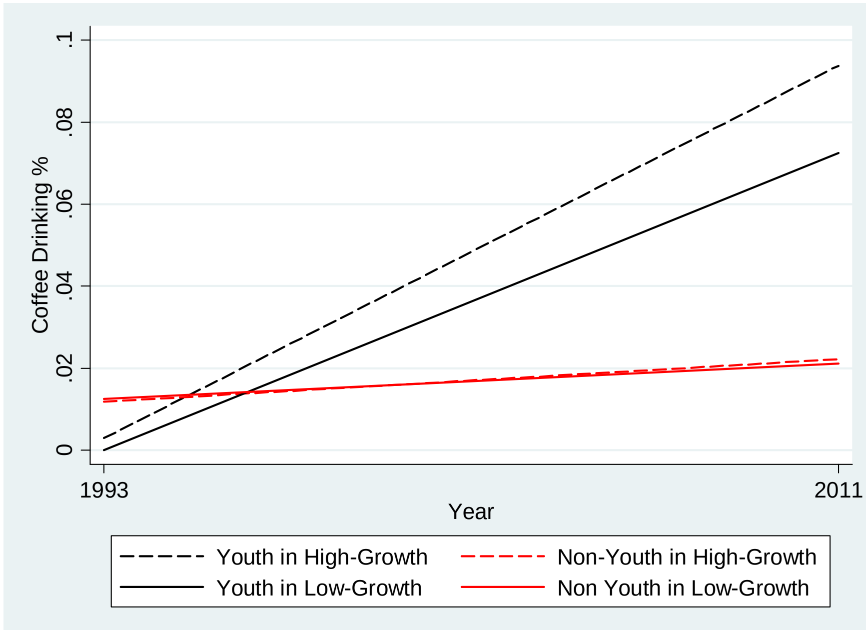
(b) Tea



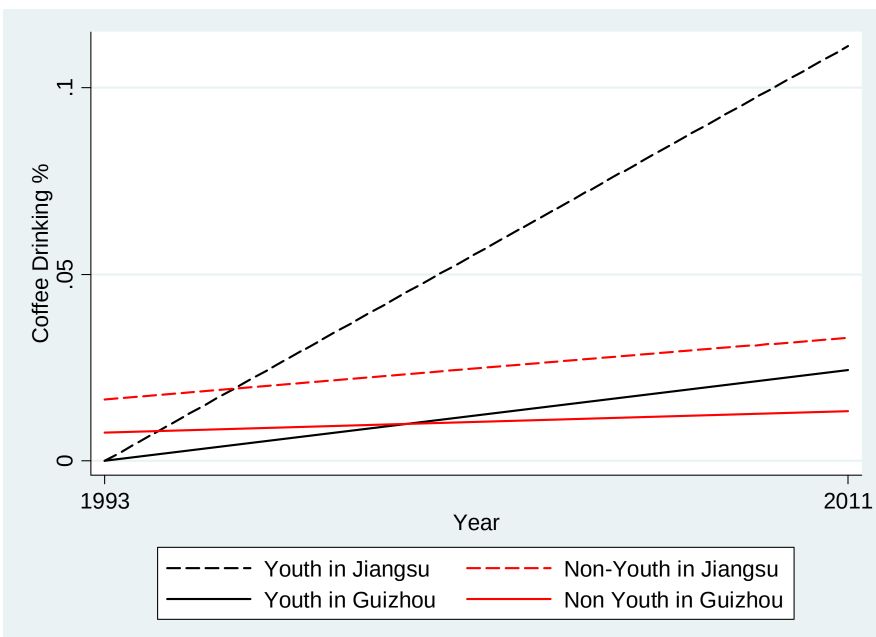
Note: The above figures show the adoption of (a) coffee and (b) tea by age. The solid line plots the raw data while the dashed line subtracts an individual's average adoption propensity from each of their observations

Figure 5: Coffee Penetration, by Provincial Income and Youth Demographic

(a) Province Classification by Above and Below Median Income



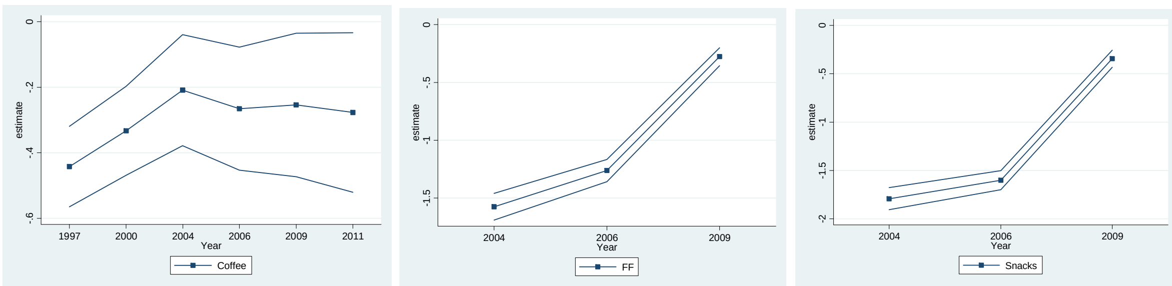
(**b**) Province Classification by Richest (Jiangsu) and Poorest (Guizhou) Province



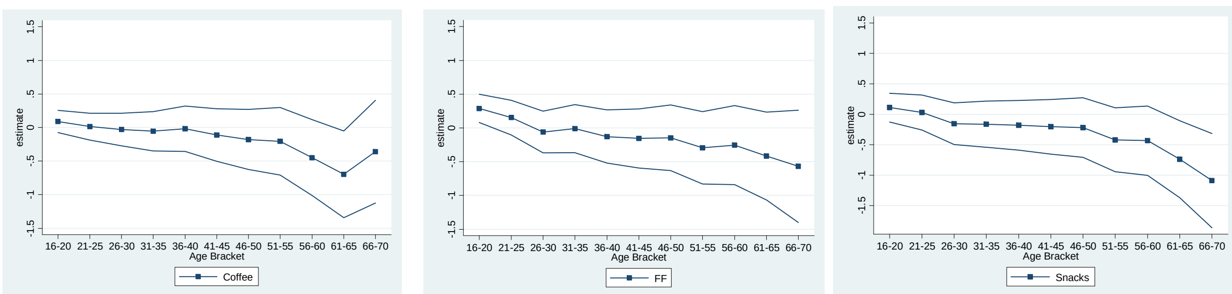
Note: (a) plots % of coffee drinkers in 1993 and 2011 by demographics (youth/non-youth) and provincial economic prosperity (higher than median income in sample/lower than median income sample (b) plots % of coffee drinkers in 1993 and 2011 by demographics (youth/non-youth) and richest (Jiangsu)/poorest (Guizhou) province

Figure 6: Cohort, Age and Time Fixed Effects from Table 2

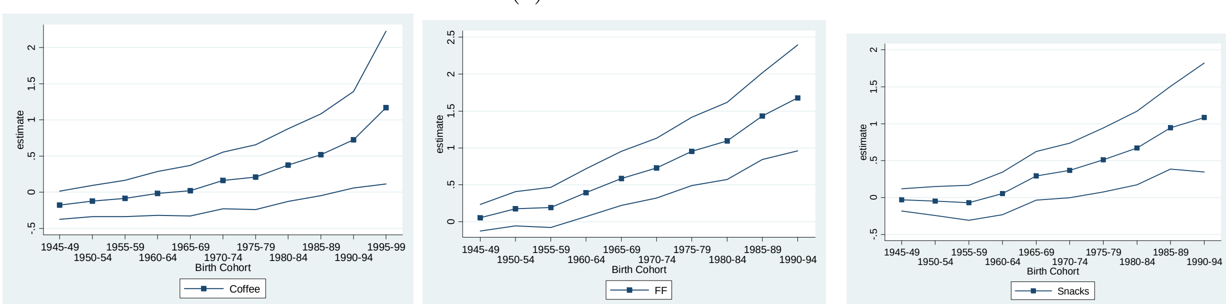
(a) Period/Year Effects



(b) Age Effects

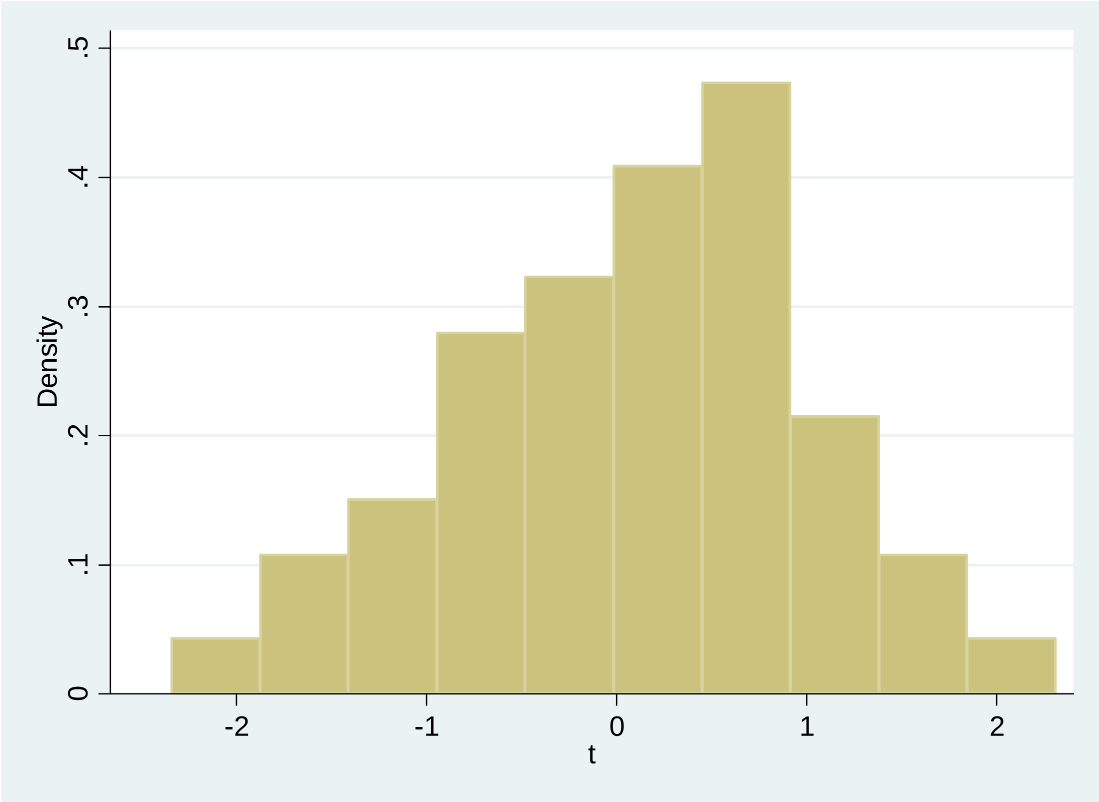


(c)Cohort Effects



Note: The figures above plot coefficients from specification (1.1) in Table 2 for all three modern categories. (a), (b) and (c) plots the estimated coefficients from the five-year age bracket dummies, birth cohort coefficients and survey year dummies respectively. Standard error bars are displayed.

Figure 7: Distribution of t-statistics in Placebo Tests



Notes: The figure plots the distribution of the t-statistic of the coefficient estimate for the baseline coffee regression in (1.2) runs 100 times. In each of the runs, individual is assigned a random false birth year.

Table 1: Summary Statistics

A. Key Variables

	N	Mean	SD	Min	Max
HH Income	46546	10733	18813	300	657000
Province GDP per capita	44044	15375	13736	1234	62290
Age	46509	42.788	12.424	6	71
Male %	46546	0.499	0.500	0	1
Rural %	46546	0.672	0.470	0	1
Student %	44898	0.044	0.205	0	1

B. Overall and Cohort-wise Category Choice

	Mean	1940s	1950s	1960s	1970s	1980s	1990s
Coffee Drinking	0.025	0.009	0.018	0.025	0.037	0.070	0.097
Coffee Share	0.031	0.007	0.018	0.027	0.048	0.153	0.255
Tea Drinking	0.389	0.426	0.409	0.397	0.368	0.262	0.207
Fast Food	0.497	0.145	0.215	0.342	0.516	0.733	0.865
Snacks	0.355	0.308	0.384	0.503	0.638	0.811	0.931
Fruit (preference)	3.650	3.563	3.624	3.683	3.707	3.768	3.979
Vegetables (preference)	3.787	3.806	3.838	3.824	3.790	3.699	3.697

Table 2: **Cohort Effects in Non-Traditional Category Consumption**

	(1)	(2)	(3)
	Coffee	**Fast Food**	**Snacks**
1945-49	-0.168	0.0537	-0.0305
	(0.12)	(0.11)	(0.09)
1950-54	-1.25E-01	1.76E-01	-4.60E-02
	(0.13)	(0.14)	(0.12)
1955-59	-0.0864	0.193	-0.07
	(0.15)	(0.17)	(0.15)
1960-64	-0.0183	0.394**	0.0548
	(0.18)	(0.20)	(0.18)
1965-69	0.0193	0.587***	0.294
	(0.21)	(0.22)	(0.20)
1970-74	0.164	0.727***	0.368
	(0.24)	(0.25)	(0.23)
1975-79	0.207	0.954***	0.511*
	(0.27)	(0.28)	(0.26)
1980-84	0.37	1.094***	0.671**
	(0.30)	(0.32)	(0.30)
1984-89	0.514	1.432***	0.946***
	(0.34)	(0.36)	(0.34)
1990-94	0.778*	1.679***	1.084**
	(0.40)	(0.44)	(0.45)
1995-99	1.162*		-0.479
	(0.64)		(0.83)
log Current Income	0.246***	0.118***	0.106***
	(0.02)	(0.01)	(0.01)
Time, geog, agebr FE	yes	yes	yes
Observations	46,507	12,604	12,422

Note: The specification in all columns is a probit where the dependent variable is a dummy =1 if the individual consumes the category in the row heading and 0 otherwise. The sample for coffee is larger as the variable is available starting from 1993 but only from 2004 for the other categories. All columns control for year, province, rural/urban and five-year age bracket fixed effects. The omitted cohort is those born in 1940-44. The observations for fast food in the 1995-99 cohort bin were collinear (no variation in adoption) and thus dropped. Robust standard errors clustered at the individual level are in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: "Prosperity" Effect in Non-Traditional Category Consumption

(a) Coffee

	(1)	(2)	(3)	(4)	(5)
log Income 18-28	0.201***	0.142***	0.141***	0.143***	0.145***
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
log GDPpc 18-28	0.120***	0.216***	0.200**	0.231*	0.229*
	(0.04)	(0.05)	(0.09)	(0.12)	(0.12)
log Current Income		0.170***	0.170***	0.160***	0.162***
		(0.03)	(0.03)	(0.03)	(0.03)
log Current GDPpc		0.23	0.24	0.21	0.00
		(0.40)	(0.40)	(0.40)	(0.62)
Male	0.056	0.0668	0.0674	0.0664	0.0683
	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Student	-0.0116	0.0468	0.0443	0.0557	0.0613
	(0.051)	(0.07)	(0.07)	(0.072)	(0.07)
Rural	-0.654***	-0.605***	-0.605***	-0.599***	-0.597***
	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Time & Province FE	yes	yes	yes	yes	yes
Cohort FE	no	no	yes	yes	yes
Age Bracket FE	no	no	no	yes	yes
Province time-trends	no	no	no	no	yes
Observations	17,729	11,812	11,812	11,651	11,651

(b) Fast Food

	(1)	(2)	(3)	(4)	(5)
log Income 18-28 yrs	0.196***	0.177***	0.177***	0.176***	0.175***
	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
log GDPpc 18-28 yrs	0.358***	0.378***	0.373***	0.304**	0.300**
	(0.04)	(0.05)	(0.09)	(0.12)	(0.12)
log Current Income		0.157***	0.158***	0.149***	0.153***
		(0.02)	(0.02)	(0.02)	(0.02)
log Current GDPpc		0. 195	0.191	0.258	3.210***
		(0.62)	(0.62)	(0.63)	(1.19)
Male	0.0919**	0.0644	0.0644	0.0712	0.074
	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Student	0.105*	0.0693	0.0677	0.0631	0.0589
	(0.05)	(0.08)	(0.08)	(0.08)	(0.08)
Rural	-0.727***	-0.678***	-0.678***	-0.669***	-0.676***
	(0.05)	(0.06)	(0.06)	(0.06)	(0.061)
Time & Province FE	yes	yes	yes	yes	yes
Cohort FE	no	no	yes	yes	yes
Age Bracket FE	no	no	no	yes	yes
Province time-trends	no	no	no	no	yes
Observations	3,469	3,469	3,469	3,440	3,362

(c) Western Snacks

	(1)	(2)	(3)	(4)	(5)
log Income 18-28 yrs	0.187***	0.158***	0.132***	0.137***	0.138***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
log GDPpc 18-28 yrs	0.353***	0.346***	0.365***	0.422***	0.241*
	(0.03)	(0.04)	(0.05)	(0.09)	(0.13)
log Current Income			0.126***	0.130***	0.130***
			(0.02)	(0.03)	(0.03)
log Current GDPpc			1.600**	1.637**	1.03
			(0.80)	(0.80)	(1.27)
Male		0.256***	0.236***	0.227***	0.244***
		(0.05)	(0.06)	(0.06)	(0.06)
Student		0.136**	0.115	0.109	0.0719
		(0.06)	(0.08)	(0.08)	(0.09)
Rural		-0.501***	-0.458***	-0.462***	-0.452***
		(0.05)	(0.07)	(0.07)	(0.07)
Time & Province FE	yes	yes	yes	yes	yes
Age Bracket FE	no	no	no	no	yes
Cohort FE	no	no	no	yes	yes
Province time-trends	no	no	no	yes	yes
Observations	3,392	5,641	5,459	3,392	3,362

Note: The specification in all panels (a)-(c) is a probit where the dependent variable is a dummy =1 if the individual consumes the category and 0 otherwise. The columns progressively add covariates and the FE indicated. The last column in all tables controls for a province-wise linearly increasing time-trend. All columns control for year as well as relationship to the head. Robust standard errors clustered at the individual level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

(d) PIY Explanatory Power: Pseudo- R^2

	Demographics	Direct PIY Effect	Indirect PIY Effect	Cohort FE	Age Bracket FE
Coffee	0.151	0.169	0.173	0.173	0.174
Fast Foods	0.328	0.355	0.368	0.368	0.37
Snacks	0.31	0.331	0.341	0.342	0.345

(e) PIY Effect on Average Consumption

	Direct PIY	Indirect PIY	Direct+Indirect PIY
Coffee	0.12	0.18	0.28
Fast Foods	0.06	0.11	0.17
Snacks	0..05	0..09	0.13

Table 4: "PIY" Effect—Other Measures

(a) Coffee Share

	(1)	(2)	(3)	(4)
log Income 18-28	0.171***	0.116**	0.118**	0.123***
	(0.0459)	-0.0469	(0.05)	(0.05)
log GDPpc 18-28 yrs	0.305***	0.373***	0.387***	0.437***
	(0.0624)	(0.063)	(0.11)	(0.15)
log Current Income		0.115***	0.116***	0.119***
		(0.04)	(0.04)	(0.04)
log Current GDPpc		0.16	0.16	0.12
		(0.47)	(0.47)	(0.47)
Male	0.387***	0.413***	0.412***	0.403***
	(0.07)	(0.07)	(0.07)	(0.07)
Student	0.0295	0.0386	0.0387	0.0381
	(0.086)	(0.09)	(0.09)	-0.0874
Rural	-0.426***	-0.405***	-0.407***	-0.405***
	(0.07)	(0.07)	(0.07)	(0.07)
Time & Province FE	yes	yes	yes	Yes
Cohort FE	no	no	yes	Yes
Age Bracket FE	no	no	no	Yes
Observations	3,907	3,907	3,907	3,907

(b) Preferences

	Snacks		Fast Food	
	Direct	Both	Direct	Both
log HH income 18-28 yrs	0.0717***	0.0602**	0.0987***	0.0906***
	(0.02)	(0.02)	(0.03)	(0.03)
log GDPpc 18-28 yrs		0.214***		0.149*
		(0.08)		(0.08)
log Current Income	0.0416**	0.0433**	0.0680***	0.0693***
	(0.02)	(0.02)	(0.02)	(0.02)
log Current GDPpc	0.696*	0.689*	0.786**	0.775**
	(0.37)	(0.37)	(0.39)	(0.39)
Male	0.238***	0.227***	0.107**	0.0996**
	(0.04)	(0.04)	(0.04)	(0.04)
Student	0.0216	0.0259	-0.00961	-0.00589
	(0.052)	(0.052)	(0.054)	(0.054)
Rural	-0.191***	-0.189***	-0.325***	-0.323***
	(0.04)	(0.04)	(0.04)	(0.04)
Time, geo, agebr,cohort FE	yes	yes	yes	yes
Observations	6,028	6,028	6,025	6,025

Note: Panel (a) reports tobit regressions with share of coffee, relative to coffee and tea as the DV. Panel (b) reports ordered logit regressions with stated preference for the indicated category on a scale of 0-4 as the DV. All columns control for year, province, rural/urban, five-year age bracket and cohort fixed effects as well as relationship to the head. Robust standard errors clustered at the individual level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 5: "Prosperity" Effect in Non-Traditional Category Consumption: Accounting for Province Level Unobservable Trends

	(1)	(2)	(3)
	Coffee	**Snacks**	**Fast Food**
log HH income 18-28 yrs	0.145***	0.138***	0.175***
	(0.04)	(0.04)	(0.04)
log GSPpc 18-28 yrs	0.229*	0.241*	0.300**
	(0.12)	(0.13)	(0.12)
log Current HH Income	0.162***	0.130***	0.153***
	(0.03)	(0.03)	(0.02)
log Current GSPpc	(0.00)	1.03	3.210***
	(0.62)	(1.27)	(1.19)
Modernity			
Gender	0.0683	0.244***	0.074
	(0.06)	(0.06)	(0.06)
Student	0.0613	0.0719	0.0589
	(0.07)	(0.09)	(0.08)
Rural	-0.597***	-0.452***	-0.676***
	(0.06)	(0.07)	(0.067)
Province time-trends	yes	yes	yes
Time, geo,agebr,cohort FE	yes	yes	yes
Observations	11,651	3,362	3,440

Notes: Columns 1-3 control for a linear trend at the province-level.

Table 6: Falsification Test—PIY Effect in Traditional Category Adoption

	Fruits	Vegetables
	Preference	Preference
log HH income 18-28 yrs	0.0396	0.0875
	(0.04)	(0.05)
log GDPpc 18-28 yrs	0.0837	-0.209
	(0.13)	(0.16)
log Current HH Income	0.0639**	0.02
	(0.03)	(0.03)
log Current GDPpc	0.52	-0.49
	(0.61)	(0.73)
Male	0.435***	0.147**
	(0.07)	(0.07)
Student	0.0239	0.111
	(0.084)	(0.094)
Rural	0.202***	0.147**
	(0.06)	(0.07)
Time, geo,agebr,cohort FE	yes	yes
Observations	4,849	5,700

Notes: Ordered logit regressions with stated preference for the indicated category on a scale of 0-4 as the DV

Table 7: Aspirations: Heterogeneity by Income Bracket in Coffee Consumption

	(1)	(2)	(3)	(4)
	Low Relative Income as Youth		**Low Relative Income as Adult**	
	Adoption	Share	Adoption	Share
Low Relative Income	-1.553*	-2.318**	-0.845	-1.365*
	(0.81)	(1.05)	(0.54)	(0.74)
Low Relative Income*log GDPpc	0.168**	0.257**	0.0613	0.116
	(0.08)	(0.11)	(0.05)	(0.07)
log HH income 18-28	0.144***	0.138**		
	(0.05)	(0.06)		
log GDPpc 18-28	0.188	0.360**		
	(0.13)	(0.16)		
log Current Income	0.157***	0.129***	0.200***	0.198***
	(0.03)	(0.04)	(0.02)	(0.03)
log Current GDPpc	0.11	0.10	0.615***	0.39
	(0.41)	(0.48)	(0.23)	(0.29)
Male	0.0874	0.410***	-0.0175	0.293***
	(0.06)	(0.07)	(0.04)	(0.04)
Student	0.055	0.0524	0.0812	0.0449
	(0.07)	(0.09)	(0.06)	(0.08)
Urban	-0.569***	-0.392***	-0.569***	-0.432***
	(0.06)	(0.07)	(0.03)	(0.04)
Time, geo,agebr,cohort FE, other covar	yes	yes	yes	yes
Observations	11,652	3,922	11,652	3,922

Note: The specification in all columns 1 and 3 is a probit where the dependent variable is a dummy =1 if the individual consumes coffee and 0 otherwise. The specification in all columns 2 and 4is a tobit where the dependent variable is the total share of coffee—ratio cups of coffee to total cups of coffee and tea. Low Relative Income at youth (adult) is defined as a 1 if youth income at age 18-28 (adult income is average over ages 30-60) is less than the mean province income. All columns control for year, province, rural/urban, five-year age bracket and cohort fixed effects as well as all the demographics from previous tables. Robust standard errors clustered at the individual level are in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: Effect of Supply Conditions on Fast Food Consumption

	Adopt	Pref
log Income 18-28	0.0940**	0.0188
	(0.04)	(0.04)
log GDPpc 18-28	0.376**	0.359***
	(0.16)	(0.14)
log Current Income	0.210***	0.181***
	(0.03)	(0.03)
log Current GDPpc	1.52	2.33
	(1.88)	(1.61)
log FF outlets 18-28	-0.117	-0.0795
	(0.09)	(0.07)
log Current FF outlets	1.958***	1.043**
	(0.598)	(0.523)
Time, geo, agebr, cohort FE	yes	yes
Observations	2,061	2,916

Note: The specification in column (1) is a probit where the dependent variable is a dummy =1 if the individual consumes the category in the row heading and 0 otherwise. In column (2) it is an ordered logit where the dependent variable is the stated preference for the indicated category on a scale of 0-4. Each column includes variables that measure the log number of fast food (KFC and McDonald's) outlets in the province during the individuals youth as well as the log of the current number of these outlets. All columns control for year, province, rural/urban, five-year age bracket and cohort fixed effects as well as Male, student status, rural/urban location and relationship to the head. Robust standard errors clustered at the individual level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1