ACCESS TO MARKETS AND RURAL POVERTY: EVIDENCE FROM HOUSEHOLD CONSUMPTION IN CHINA

M. Shahe Emran and Zhaoyang Hou*

Abstract—This paper presents evidence on the effects of access to domestic and international markets on per capita consumption of households using data from rural China. The econometric analysis uses alternative identification schemes to address the potential endogeneity of access to markets. We use straight-line distances to coastline and navigable river, along with the topography of the intervening counties, as sources of exogeneous variations. We also use identification through heteroskedasticity, which does not rely on standard exclusion restrictions. The results from alternative identification schemes show that better access to both domestic and international markets has positive effects on per capita consumption, the domestic market effect is significantly larger in magnitude, and there is complementarity between the access to domestic and international markets.

I. Introduction

IMPROVING access to markets, especially for the rural poor, is a central element of the current thinking on poverty alleviation and economic development (see World Development Reports, 2001, 2006; Sachs, 2005; Smith, 2005). There is also a growing literature on the spatial aspects of poverty and inequality in developing countries (see Kanbur & Venables, 2005; Henderson, Shalizi, & Venables, 2001). The literature has, however, focused on either access to the domestic market centers using household data or access to world trading centers using cross-country data. The existing literature thus largely ignores the roles played by both domestic and international markets in the determination of any given economic outcome, especially at the household level. Using household survey data from rural China in 1995, we provide robust evidence that better access to domestic and international markets confers substantial benefits on rural households, the effects of domestic market access being significantly larger. Also, there is evidence of complementarity between the access to domestic and international markets; a reduction in the distance to one market (say, the domestic market) yields a higher benefit for a household when it is located closer to the other market.[1]

Rural China is especially suitable as a case study to uncover the effects of market access (or lack of it) at the household level because of the Hukou system, which has restricted the geographic mobility of rural households since 1951.[2] This implies that the challenges that arise from mobility and spatial sorting of households for identification of the effects of distances to the markets is much less of a concern in rural China. However, even if household spatial sorting is not a major problem, there are other sources of potential bias in the estimated effects of market access, including possible nonrandom placement of transport infrastructure and the endogeneous emergence of markets in geographic space. We use two alternative identification schemes to address the potential endogeneity of access to markets: instrumental variables based on geography and identification based on variations in the second moment as developed by Lewbel (2012). For the instrumental variables approach, the sources of exogeneous variations in the access to markets are the fundamental geographic features correlated with the distances to the markets.[3] The fundamental geographic feature we rely on for access to international markets (the nearest seaport) is the nearest coastline from the village; for access to a domestic market center (nearest business center), it is the nearest navigable river. The location of the coastline and the navigable river are clearly exogeneous. We use straight-line distances from a village to the nearest coastline and navigable river along with topographic features of the intervening counties (for example, elevation and slope) as the instruments for identifying the effects of access to domestic and international markets.[4] We also use a recent approach developed by Lewbel (2012) that exploits heteroskedasticity for identification and does not rely on standard exclusion restrictions.[5] The

Received for publication June 16, 2008. Revision accepted for publication February 8, 2011.

* Emran: IPD and Columbia University; Hou: China Economic Research Center and Stockholm School of Economics.

We thank an anonymous referee for very helpful comments on the identification strategy used in this paper. We are grateful to Jeffrey Wooldridge, Arthur Lewbel, Steve Pischke, Martin Ravallion, Dani Rodrik, Albert Park, John Giles, Alan de Brauw, Bryan Boulier, Robert Phillips, David Ribar, Don Parsons, Stephen Smith, Bob Trost, Tony Yezer, Wally Mullin, Maggie Chen, Francisco Ferreira, Bob Goldfarb, Mushfiq Mobarak, Atonu Rabbani, Arif Mamun, Forhad Shilpi, Margriet Caswell, Michael Jerzmanowski, Bichaka Fayissa, Daniel Westbrook, and the participants in the Trade and Development Workshop at George Washington University, NEUDC conference 2006 at Cornell, and the Midwest Economic Association Annual Conference, 2006 at Chicago for useful discussions or comments on earlier versions of the paper. Special thanks to Loren Brandt for providing us with the spatial cost-of-living index in China, to Albert Park for data on county-level characteristics, and to Uwe Deichmann and Hunjong Yoo for help with calculating the distance measures used in this paper. The standard disclaimers apply.

A supplemental appendix is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00354.

[1] As explained in greater detail later in the paper, we identify the nearest business center as the relevant domestic market and the nearest seaport as the access point to the world market.

[2] A household registry system and its use in various forms of social control goes back to the days of the Xia dynasty (2070 BC–1600 BC). The Hukou system was introduced after the Communist revolution to put controls on internal migration. We discuss the implications of Hukou and its gradual relaxation for our analysis later in the paper.

[3] We are grateful to an anonymous referee for suggesting this identification strategy. An earlier working paper version used the topographic features of the intervening counties along the straight line connecting the village with the seaport and the relevant business center as identifying instruments drawing on the transport engineering literature (see Emran & Hou, 2008).

[4] For a recent paper that uses topographic features for identifying the effects of infrastructure on economic outcomes when infrastructure placement may be endogeneous, see Duflo and Pande (2007).

[5] For recent applications of heteroskedasticity-based identification, see Emran and Shilpi (2012), Mallick (2012), Emran, Robano, and Smith (2011), Emran and Sun (2011).

The Review of Economics and Statistics, May 2013, 95(2): 682–697
© 2013 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

results from these alternative identification schemes are similar and provide us with robust evidence on the effects of access to domestic and international markets.

According to the estimates reported later in the paper, a 10 kilometer reduction in the distance to a domestic market increases per capita consumption by 42.10 yuan when the household is initially located 150 kilometers away from both the domestic and international markets, averaging over the estimates from geography-based IVs and Lewbel's (2012) two-stage estimator. The corresponding increase in per capita consumption for a reduction of 10 kilometers in the distance to international market is 29.68 yuan (again averaging over the estimates from two different identification schemes). When expressed as a percentage of the 1995 rural poverty line income in China (National Bureau of Statistics estimate), they amount to 7.94% (domestic market effect) and 5.6% (international market effect). The benefits from the reduction in the distances by 10 kilometers in case of a household located initially 200 kilometers away from both domestic and international markets as a percentage of the rural poverty line income are 5.76% (domestic) and 4.02% (international) of the 1995 rural poverty line. To the best of our knowledge, this is the first paper to provide evidence on the effects of both domestic and international markets on the consumption of rural households in a developing country.

The rest of the paper is organized as follows. Section II provides a concise discussion of the related literature and thus puts the contributions of the paper in perspective. Section III discusses the conceptual framework that underlies the empirical work. Section IV provides a discussion of the econometric issues in identification and estimation of the effects of market access on household consumption. Section V is devoted to the alternative identification schemes used in this paper. Section VI describes the data sources and the main variables used in the empirical analysis. Section VII reports the results from the empirical analysis. The paper concludes with a summary of the main results.

II. Related Literature

The roles of domestic and international markets in economic development have long been at the center of debate in development economics. There is a large literature, from Rosenstein-Rodan (1943) to Murphy, Shleifer, and Vishny (1989), that emphasizes the role played by a large domestic market in initiating industrialization. The focus of a second thread of the literature has been on the advantages of an open trade regime and better access to international markets (Little, Scitovosky, & Scott, 1970; Bhagwati, 1978; Krueger, 1983). A long tradition in economics based on comparative advantage and economies of scale points to the benefits of opening up an economy to the international market. The more recent literature at the intersection of economic geography and development analyzes the effects of market access using either household or cross-country data. A number of household-level studies provide estimates of the costs of being located in a geographically remote village, away from the urban centers (Jacoby, 2000; Minten & Kyle, 1999). There is also a burgeoning literature that focuses on the effects of trade liberalization on poverty and inequality (Winters, McCulloch, & Mckay, 2004; World Development Report, 2006). The recent literature emphasizes the heterogeneity in the effects of trade liberalization (see Topolova, 2005, in the context of India). One element of this heterogeneity at the household level is its geographic location; trade liberalization might have little impact on households that are geographically remote from access points to international markets like seaports (for evidence in the context of Mexico, see Nicita, 2009). The cross-country literature, on the other hand, is concerned with the effects of a lack of access to international markets due to either geographic reasons or trade restrictions, on the per capita income of a country (Gallup, Sachs, & Mellinger, 1999; Frankel & Romer, 1999; Redding & Venables, 2004; Redding & Sturm, 2008). A small literature using cross-country data estimates the effects of the size of the initial domestic market and the degree of trade openness on the level of per capita income (Ades & Glaeser, 1999; Alcala & Ciccone, 2003). In an interesting paper, Hanson (1997) estimates the regional relative wage gradient in Mexico with respect to the distances from Mexico City and the U.S. border.

Our analysis is obviously related to the literature on the effects of reform and openness on poverty and inequality in China. There is convincing evidence that the openness to international trade and domestic market reform initiated in 1978 reduced poverty in China significantly (World Development Report, 2006; Ravallion & Chen, 2007). However, the focus of the recent debate has been on the rising inequality during the reform period (World Bank, 1997; Benjamin, Brandt, & Giles, 2005). The early contributions on the regional inequality in China were primarily concerned with the differential performance of coastal versus interior provinces (Jian, Sachs, & Warner, 1996; Chen & Fleisher, 1996). Jalan and Ravallion (2002) provide evidence of possible geographic poverty traps using household panel data. In a number of papers Benjamin et al. (2005, 2008), use Research Center for Rural Economy (RCRE) household panel data set and show that the importance of geographic location in explaining the rural inequality has declined over time. The recent literature based on household data also shows that in rural China, income inequality is driven by uneven development of nonfarm opportunities, declining agricultural prices, and differences in wage rates (Rozelle, 1994; Benjamin et al., 2005; Khan & Riskin, 1999). Benjamin et al. (2002) highlight the role of human capital along with the location of a household as determinants of poverty and inequality. A number of recent papers have analyzed the role of the Hukou system in the emergence and persistence of rural-urban income inequality (Whalley & Zhang, 2007; Au & Henderson, 2006a, 2006b). Using a simulation model, Whalley and Zhang (2007) find that most of the rural-urban inequality can be attributed to that system. In an analysis of the time series data for 1952 to 1999, Kanbur and Zhang

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

(2005) find that regional inequality in China can be explained by three key variables: the degree of openness (proxied by the average tariff rate), the degree of centralization, and the ratio of heavy industry to gross output value. As mentioned before, the literature on poverty and inequality in China has focused primarily on the role played by the access to international markets; the relative importance of local and global markets remains to be explored.

III. The Conceptual Framework and Estimating Equation

We outline a simple conceptual framework that highlights the role of access to markets in the determination of household consumption in rural areas. The geographic location of a household is one of the important determinants of its access to input, output, and labor markets. Better access to markets can increase household per capita consumption through higher income, given initial endowments, and a better consumption choice set given the income due to favorable output prices and a richer array of consumer goods available. The higher income results from economies of scale in production and marketing, a finer division of labor, adoption of new technology, and better labor market opportunities. Better market access for a household may be associated with lower input prices, higher output prices, and higher wage levels. The prices a household pays for consumer goods also depend on its location; households in a village closer to an urban center are likely to pay lower prices for consumer goods (manufacturing and services) because of lower transport costs. Close proximity to an urban center or seaport, however, also means that the general cost of living may be higher, especially because of higher housing prices and higher prices of agricultural products.[6] We use the recent spatial living cost index constructed by Brandt and Holz (2006) to take into account such heterogeneity in consumer prices.

We allow for potential interaction effects between the domestic and international market access for a given household. On a priori grounds, the relationship between the domestic and international markets can be either complementary or substitute. Access to any one market can be complementary to the access to the other when there is learning by doing and increasing returns. Access to a large domestic market allows entrepreneurs to reap economies of scale in nontraded intermediate inputs and also to use it as an effective testing ground. This facilitates learning by doing and helps them become internationally competitive. There can also be complementarity between the domestic and international markets because the domestic market center may act as a conduit for the international market. Two markets can be substitutes for each other for the simple reason of additivity constraint: if a farm sells more to the international market, it necessarily sells less to the domestic market given a certain amount of production. To put it another way, if a producer has access to the international market, the domestic market may not be an important source of demand for its product.

The estimating equation for per capita consumption at the household level can be written as

$$C_i = \beta_0 + \beta_p + \beta_1 A_i^d + \beta_2 A_i^g + \beta_3 \left(A_i^d \times A_i^g\right) + X_i^{'}\theta_i + X_i^{v'}\theta_v + X_i^{c'}\theta_c + \varepsilon_i, \quad (1)$$

where A_i^d and A_i^g are the indicators of access to domestic (superscript d) and international (superscript g) markets for household i, X_i is a vector of relevant household characteristics that determine consumption choices, X_i^v is a vector of village-level characteristics, X_i^c is a vector of county-level controls that capture differences in economic endowments across different counties, and β_p is a province-level fixed effect. The province-level fixed effects take care of any time-invariant heterogeneity in economic endowments across different provinces, and thus the estimated effects of the market access variables cannot be due to the much-discussed provincial-level differences (coastal versus interior) in economic growth in China (Jian et al., 1996; Yang, 2002). The candidates for household-level controls include resource endowments (physical assets like land and fixed productive assets, and human capital, among others), demographic composition, and also the political capital of a household as measured by membership in the Communist party (a party official or party cadre). It has been argued by a number of authors that the economic opportunity set in China may be partly determined by a household's political capital (Nee, 1992; Morduch & Sicular, 2000; Benjamin et al., 2002). However, more recent evidence based on data on twins shows that party affiliation reflects unobserved heterogeneity in ability, and there is no significant positive rent to party membership (Li et al., 2007). Given the evidence, party affiliation seems to be a good indicator of high ability in China, but it is clearly not exogeneous. In fact, most of the individual and household-level variables discussed above may not be appropriate controls when the interest is to estimate the total causal effects of access to markets on the household consumption (total derivative rather than partial derivative). These are what Angrist and Pischke (2008) call "bad controls" in the context of our application. For example, consider the variables capturing endowments such as fixed productive assets and human capital. The accumulation of physical and human capital is determined by the expected returns, which depend on access to markets. So some of the benefits of better access to markets are represented by the accumulation of physical and human capital. Thus, when we include them as controls in the consumption regressions, the estimated effects of market access are likely to be biased downward. One might argue that it might still be a good idea to control for such variables to guard against positive selection effects and interpret the estimates as the lower bounds on the effects of market access. In the empirical implementation, we report estimates of the

[6] For the agricultural goods traded in the international market, the prices faced by the households close to the seaport may be affected more by changes in the international prices.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

effects of market access on household consumption with and without such controls.

Similar issues arise with respect to public goods provided by the government to a village. We have information on the presence of a school and health clinic in a village and also on indicators of the scale of the public sector in a county such as employment in education, health, and other public facilities. These controls might be important for guarding against positive selection effects, as the government might have more difficulties in providing public goods to counties in remote geographic locations. In this case, the estimated effects of market access would be overestimated if we do not control for heterogeneity in public goods provision, as the positive effects of better public goods will be misrepresented as the effects of being closer to the markets.[7] Alternatively, the government might target public goods to relatively disadvantageous areas, including areas where ethnic minorities live. As we discuss later in the paper, in recent Chinese economic history, such poverty targeting seems to be important objective in the placement of transport infrastructure. The fiscal decentralization and revenue sharing between the central and provincial government implied that relatively poor provinces were cross-subsidized, especially from 1980 to 1993 (Jin, Qian, & Weingast, 2005). In this case, controlling for the public goods provision would underestimate the causal effects of market access on household outcomes. Since the direction of bias cannot be determined a priori, we exclude these variables from the main specification used in this paper. However, as a robustness check, we report the estimates of market access on household per capita consumption, including a set of variables representing heterogeneity in public goods provision across geographic space.

IV. Econometric Issues

The identification and estimation of the effects of distances to a market at the household level need to take into account three sources of potential bias: endogeneous location choice by the households, endogeneous emergence of marketplaces in response to the endowments of the surrounding villages in geographic space, and nonrandom placement of transport infrastructure because of targeting by the government.

A. Household Location Choice and the Hukou System

An important advantage of rural China as a case study to understand the effects of market access at the household level is the institutional constraints (the Hukou system) on geographic mobility of the rural households. The Hukou system, established in 1951 in China, is similar to an internal passport system.[8] A person's local "citizenship" and residence is defined for a child as a birthright by the mother's place of residence (until 1998). Legal residence in a village entitles residents to land for farming, township housing, job opportunities in rural township and village enterprieses, and access to local health facilities and schools.[9] A household that moves from its village risks losing all the benefits associated with the legal residence. Although it is not impossible to change one's Hukou (local citizenship), it has remained extremely restricted until recently (Chan, 1994, 2000).

Compared to other countries where there are no restrictions on geographic mobility, the problems that arise from sorting of households in geographic space for econometric identification are thus much less of a concern in rural China. The restrictions on location choice due to the Hukou system in China have been exploited by other researchers for tackling the identification issues in the recent literature (Jalan & Ravallion, 2002). Although the Hukou system was still in place in 1995 and migration from rural areas was restricted, there is evidence that starting from the mid-1980s, the pace of rural-urban migration (predominantly short-term migrants, the so-called floating population) increased due to gradual relaxations in the migration controls (Chan, 2000; Mallee, 1995). It is, however, important to recognize that while short-term migration has increased over time, it is very unusual for an entire household to migrate to a city or another village (Jalan & Ravallion, 2002; Mallee, 1995).[10] Moreover, better access to urban labor markets for short-term employment and the benefits (like remittances) that follow from increased temporary migration are some of the benefits of the better market access we are trying to capture in this paper.[11] Also, migrants are likely to have exposure and access to a richer set of consumer goods, which might affect the consumption of the household members residing in the villages.

B. Emergence of the Markets and the Endowments of the Surrounding Villages

Another potential difficulty in identifying the causal effects of market access on household per capita consumption is the possibility that market centers emerge in places where the surrounding villages are of exceptional economic potential. If the business centers and seaports emerged historically in a certain location to take advantage of such economic potential of the surrounding counties, then an observed positive correlation between the proximity to markets and better economic outcomes may be driven by the fact that the closer counties have better endowments. It is thus important to control for differential economic endowments across different counties in our sample. In the empirical analysis, we include a rich set of indicators of economic potential of a county to account for such heterogeneity. We control for a number of geographic

[7] Thanks to an anonymous referee for pointing out this possibility.

[8] The transport infrastructure in China in the early 1950s was very underdeveloped; more than 50% of provincial capitals had no rail connections with Beijing (Leung, 1980).

[9] An example of the reach of the Hukou is that a marriage certificate can be issued by a person's own hukou only.

[10] Thanks to John Giles for making this point.

[11] Benjamin et al. (2005) show that the remittances from temporary migrants equalize rural income distribution in China. Giles and Yoo (2007) show that precautionary savings are lower for both poor and nonpoor households in the villages with larger migrant network.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

and agroclimatic features of a county: the mean and variance of elevation and of slope, average annual rainfall, and average temperature, among others. These geographic and agroclimatic characteristics of the county of residence of the household are critical for our geography-based identification scheme. In the absence of controls for the geography and topography of the county of residence of a household, the instruments based on geography and topography of the other counties used in this paper may not satisfy the exclusion restrictions as the instruments may capture the effects of the own geography.

C. *Targeting and Endogeneous Placement of Transport Infrastructure*

The OLS regressions of per capita household consumption on distances to domestic and international markets may also fail to identify the causal effects because of endogeneous placement of road and railway infrastructure. The placement of public infrastructure like roads and railways may be motivated, for example, by the goals of poverty alleviation or ensuring access to raw materials, including energy sources, especially coal mines. If the objective of poverty alleviation is important, it is possible that the households located closer to such infrastructure are relatively disadvantaged, with systematically lower consumption levels reflecting unfavorable unobserved endowments. With such poverty targeting of the transport infrastructure, the OLS regressions will tend to underestimate the benefits from a reduction in the distance to a market. In fact, in our empirical analysis, we find that the OLS estimates are lower than the estimates based on instrumental variables and heteroskedasticity-based identification, which is consistent with such poverty-targeted geographic placement of roads and railway routes.[12] There is evidence that geographic targeting based on poverty has been an important objective for the Chinese government and colonial powers like the United Kingdom, France, and Japan and donors like the Soviet Union USSR and theAsian Development Bank (ADB) (for a discussion of poverty objectives in the context of a recent railway project in China, see ADB, 2004; for colonial China, see Leung, 1980; and for early communist China, see Wu, 1967, and Leung, 1980).[13] However, the objectives underlying the construction of railway roads in China have varied over time; the other dominant goals in Chinese railroad history include defense imperatives and connecting manufacturing belts with energy sources, especially coal mines. For example, "During the republic era (1911–1948), the expansion of the railway beyond the traditional Chinese ecumene into the southwest, central north, and northwest was the result of war, rather than national development priorities" (Leung, 1980, p. 89).

V. Two Approaches to Identification: IV and Heteroskedasticity

A. *Instrumental Variables-Based Identification*

To identify the causal effects of market access, we need to find a way to isolate the variations in distances to markets caused by exogeneous factors. We rely on geography as the source of such exogeneous variations. The exogeneous geographic feature used for the domestic market access is the nearest navigable river, and for the international market, it is the nearest coastline. These locations are clearly exogeneous. We use linear distances to the navigable river and the coastline, along with the topographic and agroclimatic features of the intervening counties, as instruments. It was not possible to find navigable rivers for some households in the CHIP 1995 data set used in this paper (at least part of the river is not navigable). The empirical analysis is thus based on 7,462 rural households out of a total of 7,988 households in the CHIP 1995 data set.[14]

We have data on a rich set of geographic and topographic features for the counties situated on the straight-line distances from the county of residence of a given household to the relevant navigable river and coastline.[15] The county-level topographic features include, among others, mean slope and mean elevation, maximum and minimum mean slope, maximum and minimum mean elevation, and variance (or standard deviation) of elevation and of slope. The topographic features of the counties are likely to be correlated with the actual road and rail distance to markets because of possible influence on the placement of transport infrastructure. The topography of the intervening counties might influence the choice of the transport route, as emphasized in the transport engineering literature (for a discussion of the importance of topography in the placement of railway route in general see Armstrong, 1998, and AREMA, 2006; in the specific context of China, see Emran & Hou, 2008). This is especially relevant for counties where the angle between the linear distance to the market and that to the relevant exogeneous end point (navigable river or coastline) is not too wide. The idea of using topographical features as identifying instruments in the presence of the endogeneous placement of infrastructure projects is, however, not novel in the literature. (For a recent example, see the analysis of the effects of large-scale dams construction in India by Duflo & Pande, 2007.)

[12] Another factor behind the estimated downward bias in the OLS estimates is measurement error. Since we use instrumental variables, it takes care of any such potential attenuation bias also.

[13] According to Wu (1967), during the first five-year plan period (1953–1957), more new highways were built in less developed regions than in developed ones. During this period, China invested heavily in transport network, especially in building new railroads (Comtois, 1990). See also Leung (1980) on railroads.

[14] The empirical analysis in an earlier working paper version of the paper (Emran & Hou, 2008) was based on the 7,988 households. The conclusions arrived at in that version are very similar to the ones reported here, although numerical magnitudes differ. However, the instruments used in the earlier version were based on the topographic features of the counties situated along the linear distance connecting a household to the relevant market.

[15] We thank Albert Park for providing us with the data set on county-level geographic and topographic characteristics.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

Given the large number of topographical variables that can potentially be used as instruments, along with the linear distances to the navigable river and coastline, we have to guard against the possibility of overfitting in the first stage, which can invalidate the IV estimates.[16] We thus use a subset of the topographic features as instruments and perform a robustness check by changing the set of instruments.[17] (For a complete list of the county topographic features used as instruments please see the appendix.) These county-level topography measures are weighted by the segment of the straight-line distance contained within a given county; for example, if the mean elevation is the same in two counties, the weighted mean elevation is higher when the county is larger, and thus the line segment is longer.[18]

To ensure the validity of the exclusion restrictions, we control for a set of geographic and topographic features, including mean and variance of both elevation and slope of the county of residence of a household.[19] A related objection to the exclusion restrictions imposed on the topographic features of the counties arrayed on the straight-line distances is that they might proxy for the economic potential of those intervening counties. To allay this concern, we control for the average gross social product of the other counties along the linear distance as an additional control. The topographic features of the counties along the linear distances to the coastline and navigable river used as instruments thus cannot proxy for the income levels of those counties.

B. *Heteroskedasticity-Based Identification*

We also use an alternative approach to identification that does not rely on standard exclusion restrictions that was developed by Klein and Vella (2010), and Lewbel (2012). In particular, we implement Lewbel's (2012) two-stage estimator, which exploits heteroskedasticity for identification. As Lewbel showed, identification can be achieved without imposing the standard exclusion restrictions when there exist some exogeneous variables in the structural equation and errors are heteroskedastic. We provide a brief discussion of the Lewbel two-stage estimator in section VIIB.

[16] When we use all 34 of the topographic features available as instruments, the centered R^2 is 1 in the first-stage regression for international distance. This implies that the predicted value from the first stage is virtually identical to the original endogeneous variable, thus invalidating the IV estimates. We are grateful to Jeffrey Wooldridge and Steve Pischke for clarifications on the hazards of overfitting in the IV regressions.

[17] If we use principal components to reduce the dimension of the instruments following the advice of Amemia (1966) and Bai and Ng (2008), the conclusions reported in this paper remain intact. Results based on principal components are available from the authors.

[18] We thank an anonymous referee for suggesting the straight-line segments as weights.

[19] The use of geographic features as instruments for access to global markets in the cross-country regressions has been subject to criticism. For example, Durlauf, Johnson, and Temple (2004) argue that the geographic features of a country are important determinants of per capita income or economic growth in their own right and are not likely to satisfy the exclusion restrictions. Our approach is immune to such criticism as we are controlling for the geographic features of the county of residence of a household.

VI. Data and Variables Description

A. *Data Sources*

The data used in this paper come from the Chinese Household Income Project 1995 (ICPSR 3012), a household survey designed by a group of international economists and the Economics Institute of the Chinese Academy of Social Sciences (CASS). We use the rural sample, selected from the parent sample of 67,340 households in 1995 chosen by the China's National Bureau of Statistics (NBS). It covers 7,998 rural households (representing 34,739 individual household members) in nineteen provinces.[20] The sample selection follows the equidistant selection method. The method is applied to the selection of counties within the province, as well as teams within the county and households within the team. The unit to be selected is ranked, usually by per capita income, and selection is made at fixed intervals after a random start. Generally for counties with a population over 450,000, the number of households selected is eighty, and for those under 450,000, the corresponding number is sixty. Since ten households are usually surveyed from each team, this implies that either eight or six teams are surveyed, depending on whether the county is above or below 450,000 in population.

The surveys were administered in the selected counties by the NBS survey team between March and June 1996 for the reference year 1995. Data for the CASS survey were compiled largely from the daily diaries and cash and goods transaction books maintained for the NBS's annual household income and expenditure survey. (For a complete description of the survey for 1995, see Eichen & Zhang, 1993, and Riskin, Zhao, & Li, 2000.) As mentioned before, it was not possible for some households to find a navigable river because at least part of the river is not navigable. We thus exclude these households, so our analysis is based on 7,462 rural households in nineteen provinces in 1995. We complement the basic household-level data with estimated travel distances by road and railroad and information on county characteristics from different sources (see the appendix).

B. *Variables Description*

Here we provide a brief discussion of the main variables used in our analysis. (For a complete list of variables, see the appendix.) Consumption is defined as the annual total consumption expenditure, excluding taxes and fees and the part of medical consumption subsidized by government or covered by insurance. The household's own consumption is valued at market price. Consumption expenditure is then deflated by the spatial price deflator of Brandt and Holz (2006).

The most important variables for our analysis are the measures of access to domestic and international markets. As

[20] The provinces are Beijing, Hebei, Shanxi, Liaoning, Jilin, Jiangsu, Zhejiang, Anhui, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Sichuan, Guizhou, Yunnan, Shaanxi, and Gansu.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

noted earlier, we use travel distance by road and railway to represent access to markets. The nearest business center represents the domestic market, and the nearest seaport is taken as the access point to the international market.[21] The domestic market centers (nearest business centers) comprise nineteen provincial capitals and fifteen other large cities with gross output value of industry and manufacturing more than Rmb 18 billion in 1995. (For a list of the business centers, see appendix II in the online supplement.) The distance to the international market is calculated by adding the rail distance from the business center to the nearest seaport. Following the recent empirical literature, we measure distance in logarithms, which captures the notion of diminishing marginal effects of distance on household outcome.[22] We thus use a semilog specification with the dependent variable per capita consumption expressed in level, although a double-log specification is standard in the literature on gravity equations in international trade. A double-log specification with log per capita consumption as the dependent variable can be interpreted as a concave utility function implying diminishing marginal utility. Although widely used for its analytical tractability, there is, to the best of our knowledge, no convincing empirical evidence in favor of this particular functional form. We provide estimates of benefits from better access to markets without imposing any arbitrary utility function. The estimated income gains due to better market access reported in this paper can be used as inputs to welfare analysis when combined with a functional form of the indirect utility function chosen by the policy analyst. Although we believe that the semilog specification is appropriate for providing inputs to policy analysis, a reader might still wonder if the qualitative conclusions are robust to alternative functional for specifications. We thus report the results from the double-log and level-level specifications, although we emphasize that one should not rely on the estimates from these alternative functional forms.

In addition to the access to a markets, the scale of economic activity in a business center or seaport is also important in determining the economic opportunities available to the surrounding villages. To isolate the effects of access to markets, we need to control for the heterogeneity in the market size (scale of economic activity) that different households face. As an indicator of the size of the domestic market, we use the GDP of the nearest business center.

China is semilandlocked, with a coast on its eastern and southeastern sides. This geographic feature determines that the access points to international trade are not diffused. Therefore, the degree of difficulty in participating in the global marketplace is to a large extent attributable to the varying distances from the major seaports on the eastern and southeastern coasts. We measure a rural household's access to international markets by the road and railway distance from the county seat to the nearest seaport, which is one of the seven seaports that handled approximately 70% of China's total cargo with the rest of the world in 1995: Dalian, Guangzhou, Ningbo, Qingdao, Qinhuangdao, Shanghai, and Tianjin. As an indicator of the scale of economic activities in different seaports, we use the volume of cargo handled by a seaport.

[21] It is standard in the literature to use the seaports as access points to the international markets. See, for example, Ades and Glaeser (1999).

[22] More specifically the distance variable is defined as log(1+distance in km).

VII. Empirical Results

A. *Results from Geography-Based Instrumental Variables Approach*

Table 1 reports the results of the estimation of the per capita consumption equation by OLS, 2SLS, and GMM estimators. The standard errors are clustered at the county level and also corrected for arbitrary form of heteroskedasticity. We report the GMM estimates for the instrumental variables approach in addition to the standard 2SLS estimates, because GMM is more efficient in the presence of heteroskedasticity (Baum, Schaffer, & Stillman, 2003). The estimates provide strong evidence that better access to both domestic and international markets has a statistically significant and quantitatively important positive effect on per capita household consumption in rural China. The evidence also shows that the effect of the domestic market is significantly larger and domestic and international market access are complementary to each other.

The first column shows the simple OLS estimates that do not correct for any potential endogeneity in the access to markets. The estimated coefficients of distances to the domestic and international markets are negative and statistically highly significant with p-values equal to 0. The domestic market effect is about 16% higher than the international market effect, not taking into account the interaction effect. The coefficient of the interaction of domestic and international markets is positive, implying complementarity between the benefits from better access to domestic and international markets; a marginal reduction in the distance to domestic market yields a higher benefit for a household if it is also located closer to the international market. The interaction effect is also precisely estimated; it is significant at 1%.

The second and third columns present the estimates from GMM and 2SLS, respectively, using sets of identifying instruments based on straight-line distances to the navigable river and coastline, along with the topographic features of the other (other than the county of residence of the household) counties located along the straight-line distances. Columns 2 and 3 use eighteen topography variables (nine each for domestic and international), plus the straight-line distances and their interactions as the instruments. (For details of the topographic features used as instruments, see the appendix.) The instruments satisfy the diagnostic tests for exogeneity comfortably. The Hansen's J statistics shows that the null of exogeneity for the set of instruments cannot be rejected at any reasonable significance level (the p-value is 0.23). The instruments also have excellent explanatory power according

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 1.—MARKET ACCESS AND PER CAPITA CONSUMPTION

The dependent variable is household per capita consumption in 1995 (yuan, RMB) deflated by the spatial CPI of Brandt and Holz (2006).

Variables	OLS	GMM-1	2SLS	GMM-2	Lewbel Two-stage Estimator
Log of distance to domestic market (km)	−893.12	−1,819.56	−2,229.45	−2,094.72	−1,813.80
	(210.95)***	(454.33)***	(606.57)***	(536.56)***	(358.99)***
	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.00]
Log of distance to international market (km)	−764.58	−1,500.48	−1,826.25	−1,686.77	−1,287.52
	(151.19)***	(304.82)***	(438.01)***	(384.75)***	(290.57)***
	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.0]	[p-value = 0.0]
Interaction of log distances	89.87	242.36	314.69	285.26	232.72
	(30.75)***	(66.93)***	(95.19)***	(83.59)***	(53.80)***
	[p-value = 0.003]	[p-value = 0.00]	[p-value = 0.001]	[p-value = 0.001]	[p-value = 0.002]
Observations	7,462	7,462	7,462	7,462	7,462
Test of equality of coefficients of domestic and international distance					
Ho: Coefficient of Log of distance to domestic market =	Chi square(1) = 3.15	Chi square(1) = 3.91	Chi square(1) = 4.13	Chi square(1) = 17.39	
Coefficient of Log of distance to international market	p-value = 0.0758	Prob > chi square = 0.0481	p-value = 0.0421	p-value = 0.00	
Diagnostic tests					
Obveridentification test: Hansen's J-statistic	21.99	21.99	16.87	9.39	
	[p-value = 0.2322]	[p-value = 0.2322]	[p-value = 0.15]	[p-value = 0.4022]	
Heteroskedasticity test: Pagan-Hall (1983) general test statistic	Chi square(52) = 324.595	Chi square(52) = 296.441	Chi square(46) = 259.646	Chi square(45) = 276.120	
	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.00]	[p-value = 0.00]	
Weak identification test (Kleibergen-Paap Wald F-Statistic):	19.600	19.60	23.41	___	
Stock-Yogo weak ID test critical values					
5% maximal IV relative bias	19.67	19.67	18.73	___	
10% maximal IV relative bias	10.63	10.63	10.33	___	
First-stage results summary					
Log of distance (km) to domestic market	Shea partial R^2	0.4523	0.4523	0.3563	0.8424
	(weak id) Angist-Pischke F	AP $F(19, 748)$ 81.98	AP $F(19, 748)$ 81.98	AP $F(13, 748)$ 31.53	AP $F(10, 85)$ 33.63
Log of distance (km) to international market	Shea partial R^2	0.3735	0.3735	0.2849	0.8652
	(weak id) Angist-Pischke F	AP $F(19, 748)$ 255.51	AP $F(19, 748)$ 255.51	AP $F(13, 748)$ 269.12	AP $F(10, 85)$ 18.06
Interaction of log distances	Shea partial R^2	0.3783	0.3783	0.2909	0.8265
	(weak id) Angist-Pischke F	AP $F(19, 748)$ 106.86	AP $F(19, 748)$ 106.86	AP $F(13, 748)$ 126.67	AP $F(22, 85)$ 143.59

Standard errors are robust to arbitrary heteroskedasticity and corrected for clustering at the county level. The independent variables in the above models are: variables of interest (log of road distances from county seat and the interaction of distance to the business center and that from business center to the seaport), controls (type of terrain, rainfall, temperature, elevation, variance of elevation, mean slope, standard deviation of slopes, total river length, soil quality, GDP of business center, cargo handled at seaport, average gross output value of society for the other counties along straight-line distance from county seat to navigable river and from business center to coastline), and controls at the province level (dummies for nineteen provinces). For a description of the variables, see the appendix. The Z vector used in the Lewbel (2012) estimator is composed of the following topography features for the intervening counties along the relevant straight-line distance: ratio of the weighted maximum and the weighted minimum of the mean elevation to navigable river, weighted mean elevation to coastline (m), weighted mean standard deviation of elevations to navigable river (m), and weighted mean standard deviation of elevations to coastline (m). For a description of the variabls, see the appendix. GMM-1 and 2SLS use eighteen topography features, in addition to three straight-line distance variables (distances to nearest navigable river and to coast line, and their interaction) as instrumental variables in a GMM and two-stage least squares regression, respectively. GMM-2 uses fifteen instrumental variables (twelve topography features and three straight-line distance variables) in a GMM regression. For a description of the variables, see appendix I. The number of observations drops from 7,998 to 7,462 in regressions due to the missing values of straight-line distance in 536 households. The test of heteroskedasticity in the first stage of Lewbel two-stage estimator (2012) are as follows:

1st-Stage Regression (Dependent Variable)	Breusch-Pagan/Cook-Weisberg test for heteroskedasticity	
Distance to domestic market (log)	Chi square(1) = 71.46,	p-value = 0.00
Distance to international market (log)	Chi square(1) = 2,496.25,	p-value = 0.00
Interaction of domestic and international distance	Chi square(1) = 210.13,	p-value = 0.00

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

to the relevant Shea's partial R^2 (the lowest value of Shea's partial $R^2 = 0.37$). Note that we rely on Shea's partial R^2 to gauge the explanatory power of the instruments in the first stage instead of the conventional F test, as the conventional F test for exclusion of the set of instruments can be misleading when there is more than one endogeneous variable (Shea, 1997; Cameron & Trivedi, 2005). In addition, we report the recently proposed Angrist and Pischke F test instead of the conventional F test for exclusion of the set of instruments; the lowest Angrist and Pischke F statistic is equal to 81.98. The results thus provide very strong evidence in favor of the conclusion that the set of instruments is strong, and we do not need to worry about weak IV issues. For completeness, we report the t tests of individual instruments in appendix III in the online supplement.[23]

The IV (GMM-1) estimates of the coefficients from column 2 in table 1 are $[-1{,}819.56, -1{,}500.48, 2{,}42.36]$ for domestic market, international market, and the interaction effect, respectively. The corresponding estimates from 2SLS in column 3 of table 1 are similar, although the numerical magnitudes of the estimated coefficients are larger: $-2,229.45$ (domestic), $-1,826.25$ (international), 314.69 (interaction). To see if the estimated effects are robust to changes in the set of topographical features used as instruments, column 4 reports results from a set of twelve topographical instruments in addition to the straight-line distances and their interactions. The estimated coefficients from GMM estimator are a bit larger in magnitude compared to the estimates in column 1. The results thus show that the estimated coefficients vary somewhat depending on the estimator or the set of instruments used. However, as we discuss later, the implied marginal effects of market access from different columns in table 1 are numerically similar (see table 2).[24]

The IV estimates based on topography and straight-line distances as identifying instruments are higher than the corresponding OLS estimates reported in table 1. This seems to confirm the concern that the OLS estimates might be biased downward due to poverty targeting of transport infrastructure and may also reflect possible attenuation bias because of measurement error.[25] For both sets of IVs, the direct effects of market access are statistically highly significant, with p-values equal to 0. The interaction term is also significant at the 1% level. Consistent with the OLS results, the domestic market has a significantly larger effect compared to the effect of international market. A formal test of the null that the domestic and international market effects (ignoring the interaction effect) are equal to each other is rejected at the 5% level for estimates from 2SLS (p-value = 0.02) and GMM-2 (p-value 0.04) and at the 10% level for the GMM-1 estimate in table 1.

B. *Results from Heteroskedasticity-Based Identification*

We report here the results from an alternative identification scheme based on heteroskedasticity. We first provide a brief intuitive discussion of the approach here. (For a more complete discussion, see Lewbel, 2012.) Identification can be achieved without imposing the standard exclusion restrictions if there is a vector of exogeneous variables Z and the errors are heteroskedastic. In the context of our application, the topographic features of the counties along the linear route are arguably exogeneous to the household consumption and thus are appropriate for the Z vector. As Lewbel (2012) emphasized, excellent candidates for the Z vector are the variables, which are clearly exogeneous and might be used as instruments in a given context. Accordingly, we use a set of variables based on the first and second moments of elevation and maximum and minimum slopes of the counties situated on the straight-line distances from the county of interest to the relevant navigable river and coastline. The Z vector comprises the following variables: average of the county-level standard deviation of elevation (weighted by the straight-line segment within the county) along the straight-line distances to a navigable river and coastline, weighted county-level mean elevation along the straight-line distance to a navigable river, and ratio of the maximum and minimum elevation (weighted by the line segment contained within the county) encountered along the linear distance to the coastline. We implement the two-stage estimator proposed in Lewbel (2012). In the first stage, each endogeneous variable is regressed on all of the control variables in the consumption equation (denoted below by vector X) along with the Z vector defined above, and the vectors of residuals $\hat{\xi} = [\hat{\xi}_d, \hat{\xi}_g, \hat{\xi}_{dg}]$ are retrieved. More specifically, we run the following first-stage regressions:

$$A_i^d = \phi_o + X'\gamma_1 + Z'\gamma_2 + \xi_d, \tag{2}$$

$$A_i^g = \alpha_o + X'\eta_1 + Z'\eta_2 + \xi_g, \tag{3}$$

$$A_i^d * A_i^g = \sigma_o + X'\upsilon_1 + Z'\upsilon_2 + \xi_{dg}. \tag{4}$$

These estimated residuals are then used to create instruments as follows:

$$\left(Z - \bar{Z}\right)' \hat{\xi}, \tag{5}$$

where $\bar{Z}$ is the mean of Z. As Lewbel (2012) showed, identification requires that the error terms in the first-stage regressions in equations (2–4) are heteroskedastic. Following Lewbel (2012), we use the Breusch-Pagan test of heteroskedasticity. The results show that the null of homoskedastic errors is clearly rejected in each case with a p-value equal to 0.01 or less. In the second stage, we use the set

[23] The first-stage results in appendix III in the online supplement show that the straight-line distances to the navigable river and coastline have the expected signs. We add the caveat that one should be careful in interpreting the signs of the first-stage coefficients because this stage is not based on any theory of the determination of the endogeneous variable under consideration (the road and railway distances to the markets) and thus cannot lay claim to any causality.

[24] The conclusions reached in this paper are robust to other alternative sets of topography-based instruments as long as the set is not too large or too small.

[25] Although the magnitudes of the effects are lower in OLS estimates, they are consistent with the qualitative conclusions reached in this paper using alternative identification schemes.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 2.—MARGINAL EFFECTS OF REDUCING THE DISTANCES TO THE MARKETS

	Domestic Market			International Market		
	Yuan	S.E.	p-value	Yuan	S.E.	p-value
Initial Distances (km) (Domestic, International)	Marginal Effects Based on IV estimation (GMM-1)					
150, 150	8.61	2.08	0.00	6.49	1.09	0.00
200, 200	6.31	1.52	0.00	4.71	0.77	0.00
300, 300	4.06	0.98	0.00	3.00	0.48	0.00
Mean distance: 213, 838	5.22	1.24	0.00	1.12	0.18	0.00
Median distance: 172, 702	6.57	1.57	0.00	1.37	0.23	0.00
Initial Distances (km) (Domestic, International)	Marginal Effects Based on IV Estimation (2SLS)					
150, 150	10.30	2.69	0.00	7.61	1.56	0.00
200, 200	7.53	1.96	0.00	5.51	1.11	0.00
300, 300	4.83	1.25	0.00	3.49	0.69	0.00
Mean distance: 213, 838	6.15	1.58	0.00	1.30	0.26	0.00
Median distance: 172, 702	7.75	1.99	0.00	1.60	0.33	0.00
Initial Distances (km) (Domestic, International)	Marginal Effects Based on IV Estimation (GMM-2)					
150, 150	9.83	2.41	0.00	7.11	1.37	0.00
200, 200	7.19	1.76	0.00	5.15	0.98	0.00
300, 300	4.63	1.12	0.00	3.27	0.61	0.00
Mean distance: 213, 838	5.92	1.42	0.00	1.22	0.23	0.00
Median distance: 172, 702	7.46	1.79	0.00	1.49	0.29	0.00
Initial Distances (km) (Domestic, International)	Marginal Effects Based on Lewbel Two-Stage Estimator					
150, 150	8.72	1.64	0.00	5.21	1.21	0.00
200, 200	6.39	1.19	0.00	3.76	0.88	0.00
300, 300	4.12	0.77	0.00	2.37	0.56	0.00
Mean distance: 213, 838	5.32	0.97	0.00	0.89	0.21	0.00
Median distance: 172, 702	6.69	1.23	0.00	1.09	0.25	0.00

of instruments in equation (5) to estimate the causal effects of market access. The last column in table 1 reports the results of estimating the effects of market access on per capita consumption in rural China using this two-stage estimator due to Lewbel (2012). The control variables are those used in column 2 of table 1. The estimated coefficients are $[-1{,}813.80, -1{,}287.52, 232.72]$ for domestic, international and interaction effect, respectively; all the coefficients are significant at 1%. As the diagnostic tests show, the instruments based on heteroskedasticity perform well; they have power in explaining the variations in the access to markets; the lowest value for Shea's partial R^2 is 0.82. They also satisfy the exogeneity condition comfortably; the null of exogeneity of the set of Lewbel IVs cannot be rejected with a p-value equal to 0.40 for Hansen's J statistic. The null of equality of the coefficients of domestic and international markets is rejected resoundingly with a p-value equal to 0.00.

The estimates of the effects of distances to the markets using Lewbel's (2012) approach reported in column 5 in table 1 are very similar to the IV (GMM) estimate based on our topography-based identification scheme reported in column 2. It is reassuring that the estimates from such different approaches to identification give very similar results.

According to the alternative IV estimates in columns 2 to 5 of table 1, better access to domestic and international markets has both a numerically and statistically significant (at 1%) positive effect on per capita consumption of a rural household. The effects of domestic market access are larger than those of international market access, not taking into account the interaction effect; a formal test of equality of domestic and international market effects is rejected comfortably at the 10% or lower levels of significance. The estimated coefficient of the interaction variable is positive and statistically significant at 1% providing strong evidence in favor of the conclusion that the domestic and international market effects are complementary.

C. *Marginal Effects and Benefits from a Better Access to Markets*

The estimated coefficients of equation (1) reported in table 1 and discussed above give us strong evidence of the positive and differential effects of domestic and international market access on per capita consumption. However, the estimated coefficients are not readily amenable to intuitive interpretation as the distances are in logarithms and there is an interaction effect in the specification. We, therefore, provide two intuitive measures of the comparative static effects of changes in the distances to the domestic and international markets: the marginal effects, and the effects of a reduction in the distance to a market by 10 kilometers.

Marginal effects of a reduction in the distances. Table 2 reports the marginal effects of a reduction in the distances to the domestic and international markets using both topography-based instruments and Lewbel heteroskedasticity-based identification of the coefficients in table 1. The marginal effects are heterogeneous and depend on the location of

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 3.—BENEFITS FROM REDUCING THE DISTANCES TO MARKETS BY 10 KILOMETERS
Increase in Per Capita Consumption from 10 km Reduction in Domestic and International Distances

Initial Distances (km) Domestic, International	Domestic Market (Yuan) Consumption Increase (Yuan)	Domestic Market (Yuan) Percent of Poverty Line (530 Yuan)	International Market (Yuan) Consumption Increase (Yuan)	International Market (Yuan) Percent of Poverty Line (530 Yuan)
Estimates based on IV estimation (GMM-1)				
150, 150	38.72	7.31%	29.16	5.50%
200, 200	28.11	5.30%	21.00	3.96%
300, 300	17.95	3.39%	13.25	2.50%
Estimates based on IV estimation (2SLS)				
150, 150	46.28	8.73%	34.20	6.45%
200, 200	33.53	6.36%	24.55	4.63%
300, 300	21.35	4.03%	15.41	2.91%
Estimates based on IV estimation (GMM-2)				
150, 150	44.16	8.33%	31.94	6.03%
200, 200	32.04	6.05%	22.95	4.33%
300, 300	20.44	3.86%	14.43	2.72%
Estimates based on Lewbel two-stage estimator				
150, 150	39.17	7.39%	23.40	4.42%
200, 200	28.48	5.37%	16.75	3.16%
300, 300	18.22	3.44%	10.47	1.98%

The increased consumption is estimated by 10 km reduction from one market while keeping the initial distance to the other market fixed. The poverty line of rural China in 1995 was 530 yuan per person (National Bureau of Statistics of China).

a household. Not surprisingly, the marginal effects decrease with distance: a reduction in the distance to any given market has a smaller effect on the household per capita consumption when the household is located farther away from the market to begin with. For example, for a household located 150 kilometers away from both domestic and international markets, the marginal effects of a reduction in the distances according to the IV estimates (averaging over GMM-1, 2SLS, and GMM-2) are 9.58 yuan (domestic) and 7.07 yuan (international). The marginal effects for a household initially located 200 kilometers from the markets are 7.01 yuan (domestic) and 5.12 yuan (international). For a household located at the median distances from both domestic and international markets, the marginal effects from the IV estimates (averaging over the GMM-1, 2SLS, and GMM-2 estimates) are 7.26 yuan (domestic) and 1.49 yuan (international).

The corresponding estimates of marginal effects from heteroskedasticity-based identification and Lewbel two-stage estimator are 8.72 yuan (domestic) and 5.21 yuan (international) for a household initially located 150 kilometers from both domestic and international markets For a household initially located 200 kilometers away, the marginal effects are 6.39 yuan (domestic) and 3.76 yuan (international). The marginal effects from Lewbel two-stage estimator are similar to the ones from IV estimates.

Benefits from a Reduction of the Distances by 10 kilometers. Table 3 reports the estimated effects on household per capita consumption when the distance to a market is reduced, ceteris paribus, by 10 kilometers starting from a few different initial locations. We express the changes in the per capita consumption as a percentage of the 1995 rural poverty line in China (530 yuan at 1995 prices; NBS estimate). This gives us a sense of the order of magnitude of the effects of better access to markets. For example, consider the average of the IV estimates (GMM-1, 2SLS, and GMM-2) for a household initially located 150 kilometers from both domestic and international markets. A reduction in the distance to domestic market by 10 kilometers increases per capita consumption by 43.05 yuan for this household, which is about 8% of the poverty-line income. The corresponding estimate for the reduction in the distance to the international market is 31.78 yuan, or about 6% of the poverty-line income. Given the diminishing marginal effect of distance, for a household located 200 kilometers from both markets, the corresponding increase in consumption is 31.23 yuan (a 10 kilometer reduction in distance to the domestic market) and 22.83 yuan (a reduction in the distance to the international market by 10 kilometers). In terms of the percentage of poverty-line income, they are 5.89% (domestic market) and 4.31% (international market).

Consistent with the results on marginal effects, the estimated benefits from a reduction in distance by 10 kilometers using the Lewbel two-stage estimator is very similar for the domestic market but smaller for the international market (table 3) and are not discussed in detail for the sake of brevity.

D. Additional Robustness Checks

The results from the topography-based IV approach and heteroskedasticity-based identification provide robust evidence on the effects of distances to the domestic and international markets on per capita consumption of rural households. Here we report additional supporting evidence from a series of robustness checks. For brevity, we report a single set of IV results based on the instrument set used in column 2 of table 1 and use GMM for estimation. The results of sensitivity checks are reported in table 4.

First, we check the sensitivity of the results with respect to the set of controls. As discussed before, variables such as

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 4.—ROBUSTNESS CHECKS: DIFFERENT SPECIFICATIONS

	Controls (GMM-1)			Functional Forms	
Variable	Include Household Controls	Include Public Goods	Include Household Controls and Public Goods	Log-Log	Level-Level
Log of distance (km) to domestic market	−1,518.78	−1,735.10	−1,454.90	−0.43	−0.00229
	(424.43)***	(474.32)***	(445.76)***	(0.13)***	(0.00032)***
Log of distance (km) to international market	−1,253.00	−1,397.93	−1,192.21	−0.35	−0.00135
	(285.18)***	(323.01)***	(304.90)***	(0.10)***	(0.00017)***
Interaction of log distances	194.84	230.43	187.84	0.05	0
	(62.52)***	(69.67)***	(65.46)***	(0.02)**	(0.00000)***
Hansen's J-statistic	25.168	22.268	25.841	6.51	9.40
	p-value = 0.1204	p-value = 0.2203	p-value = 0.1034	p-value = 0.7706	p-value = 0.8049
Log of distance (km) to domestic market	0.4628	0.4768	0.4822	0.3765	0.3393
Log of distance (km) to international market	0.3827	0.4133	0.4206	0.3223	0.8995
Interaction of log distances	0.3888	0.4149	0.4236	0.3238	0.4091

Number of Observations: 7,462. Standard errors (in parentheses) are robust to arbitrary heteroskedasticity, and corrected for clustering at the county level. Significant at *10%, **5%, and ***1%. Columns 1–3 have different control sets compared to the main regression (GMM-1) with variables listed in the appendix. Column 1 includes the following additional controls: individual level (household head's gender, age, minority status, government official, communist party membership, school years, farming occupation) and household level (number of children, number of old, total assigned lands used for farming, initial fixed productive assets, initial livestock, drinking water obtained from tap, lighting byelectricity, house built before 1989, debts borrowed from banks in 1995, industrial machinery owned by household). Column 2 includes public goods: village-level public facilities: health clinic, school; county-level public facilities: employment in public facilities, employment in public health service, employmentin education. Column 3 includes both household-level controls and the public goods. Log-log functional form uses log of household per capita consumption as dependent variable in GMM-1 regression in table 1. Level-level functional form uses level values of distances as independent variables in GMM-1 regression in table 1.

the physical and human capital of a household are not appropriate controls when the focus is on estimating the causal effects of market access. But one might still be interested to see how the estimated effects of market access change when we control for these variables. Column 1 in table 4 reports the results when we include a set of individual- and household-level variables in addition to the set of controls used in table 1. The estimated coefficients are broadly similar to the ones reported in table 1, but, consistent with a priori expectations, the numerical magnitudes of the market access variables are smaller. A formal test rejects the null that the estimated coefficients are equal to the coefficients in column 2 of table 1 at the 5% level of significance.[26] The results thus indicate that the household and individual-level endogeneous variables such as education and household fixed assets might capture part of the benefits from better market access.

The second column in table 4 reports the results from a specification where controls for public goods provided by the government to a county such as a school and health clinic are added to the set of regressors used in table 1. We also control for employment in public health, education, and other facilities as indicators of the scale of the public sector. Interestingly, the estimated coefficients of market access variables and their interaction are smaller than the results reported in column 2 of table 1. This is consistent with geographic targeting of public goods in rural China. However, a formal test is unable to reject the equality of the coefficients of the market access variables in column 2 of table 4 and column 2 of table 1 at the 10% level. More important from our perspective is the fact that the central qualitative conclusions reached earlier on the basis of table 1 remain intact.

[26] The hypothesis test is based on the standard errors from 300 bootstrap replications.

(Column 3 in table 4 shows the estimated effects of market access when both household-level controls and the public goods are added to the main specification. The estimates are smaller than those in columns 1 and 2 in table, but they are statistically different from the estimates in column 2 of table 1 at the 5% level. The estimates are consistent with the central conclusions of the paper.

The last two columns in table 4 report results from alternative functional forms: double-log and level-level functional forms instead of the semilog functional form used in table 1. The results are consistent with the qualitative conclusions in table 1.

Table 5 reports the marginal effects estimated using alternative specifications in table 4. The marginal effects reinforce the conclusions regarding potential biases in the estimates when one inappropriately includes endogeneous variables.

The results discussed above thus demonstrate that while the precise numerical magnitudes of the effects of market access may vary depending on the set of controls or the functional form used, the major conclusions reached on the basis of topography-based IVs and heteroskedasticity-based identification as reported in tables 1 to 3 are extremely robust.

VIII. Conclusion

This paper provides evidence on the effects of better access to domestic and international markets on the per capita household consumption in rural areas using data from China in 1995. Two alternative identification schemes are used to address the potential endogeneity of market access. We rely on exogeneous fundamental geographic features such as the location of coastline and navigable river and use insights from the transport engineering literature to find credible instruments for distances to the domestic and international markets.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 5.—ROBUSTNESS CHECKS: MARGINAL EFFECTS

	Marginal Effects of Reducing the Distances to the Markets Based on Different Specifications									
	Include Household Controls		Include Public Goods		Include Household Controls and Public Goods		Log-Log		Level-Level	
Initial Distances (km) (Domestic, International)	Domestic Market (Yuan)	International Market (Yuan)	Domestic Market (Yuan)	International Market (Yuan)	Domestic Market (Yuan)	International Market (Yuan)	Domestic Market (Yuan)	International Market (Yuan)	Domestic Market (Yuan)	International Market (Yuan)
150, 150	7.30	5.53	8.22	5.98	6.97	5.22	8.43	6.33	0.0023	0.0014
200, 200	5.35	4.02	6.02	4.34	5.11	3.80	4.93	3.68	0.0023	0.0014
300, 300	4.21	3.14	4.73	3.38	4.02	2.97	3.88	2.88	0.0023	0.0014
350, 350	2.92	2.16	3.28	2.32	2.79	2.04	1.37	1.01	0.0023	0.0014
700, 700	1.38	1.00	1.54	1.06	1.31	0.94	0.51	0.37	0.0023	0.0014
Mean distance: 213, 838	4.46	0.95	4.98	1.03	4.25	0.90	3.18	0.67	0.0023	0.0014
Median distance: 172, 702	5.61	1.16	6.27	1.26	5.35	1.10	3.87	0.78	0.0023	0.0014

Our second approach to identification uses heteroskedasticity and does not rely on the standard exclusion restrictions. In particular, we implement the two-stage estimator proposed by Lewbel (2012). The results from alternative identification schemes are similar and thus provide robust evidence on the effects of better access to domestic and international markets on the per capita consumption of rural households. The central conclusions from the empirical analysis can be summarized as follows: better access to both domestic and international markets has a significant positive effect on per capita consumption of rural households, the magnitude of the domestic market effect is substantially higher, and the domestic and international markets are complementary to each other in the determination of household consumption.

REFERENCES

ADB, "Report and Recommendation of the President to the Board of Directors on Proposed Loan to the People's Republic of China for the Dali-Lijiang Railway Project" (Manila: Asian Development Bank, 2004).

Ades, A., and E. Glaeser, "Evidence on Growth, Increasing Returns, and the Extent of Market," *Quarterly Journal of Economics* 114 (1999), 1025–1045.

Alcala, F., and A. Ciccone, "Trade, Extent of Market, and Economic Growth: 1960–1996," UPF working paper (2003).

Amemia, T., "On the Use of Principal Components of Independent Variables in Two-Stage Least Square Estimation," *International Economic Review* 7 (1966), 203–303.

Angrist, J., and S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton, NJ: Princeton University Press, 2008).

AREMA, *Practical Guide to Railway Engineering* (American Railway Engineering and Maintenance of the Way Association, 2006).

Armstrong, J., *The Railroad, What It Is, What It Does: The Introduction to Railroading* (Omaha, NE: Simmons-Boardman Books, 1998).

Au, C., and J. V. Henderson, "Are Chinese Cities Too Small?" *Review of Economic Studies*, 73 (2006a), 549–576.

——— "How Migration Restrictions Limit the Agglomeration and Productivity in China," *Journal of Development Economics* 80 (2006b), 350–388.

Bai, J., and S. Ng, "Selecting Instrumental Variables in a Data Rich Environment," *Journal of Time Series Econometrics* 1 (2008), 1577–1606.

Baum, C., M. Schaffer, and S. Stillman, "Instrumental Variables and GMM: Estimation and Testing," *Stata Journal* 3 (2003), 1–31.

Benjamin, D., L. Brandt, and J. Giles, "The Evolution of Income Inequality in Rural China," *Economic Development and Cultural Change* 53 (2005), 769–824.

Benjamin, D., L. Brandt, J. Giles, and S. Wang, "Income Inequality in the Transition Period," in L. Brandt and T. Rawski (eds.): *China's Great Transformation* (Cambridge: Cambridge University Press, 2008).

Benjamin, D., L. Brandt, P. Glewwe, and Guo Li, "Markets, Human Capital and Income Inequality in an Economy in Transition: The Case of Rural China," Richard Freeman (ed.), in *Inequality Around the World* (London: Palgrave-Macmillan, 2002).

Bhagwati, J., *Anatomy and Consequences of Exchange Control Regimes* (Cambridge, MA: Ballinger, 1978).

Brandt, L., and C. Holz, "Spatial Price Differences in China: Estimates and Implications," *Economic Development and Cultural Change* 55 (2006), 43–86.

Cameron, C., and P. Trivedi, *Microeconometrics: Methods and Applications* (Cambridge: Cambridge University Press, 2005).

Chan, K., "Urbanization and Rural-Urban Migration in China since 1982: A New Baseline," *Modern China* 20 (1994), 243–281.

——— "Internal Migration in China: Trends, Development and Scenarios," in workshop on China's Urbanization Strategy: Opportunities, Issues (&) Policy Options," Washington, DC: World Bank and PRC State Development Planning Commission (2000).

Chen, J., and B. Fleisher, "Regional Income Inequality and Economic Growth in China," *Journal of Comparative Economics* 22 (1996), 141–164.

Comtois, C., "Transport and Territorial Development in China 1949–1985," *Modern Asian Studies* (1990), 777–818.

Duflo, E., and R. Pande, "Dams," *Quarterly Journal of Economics* 122 (2007), 601–646.

Durlauf, S., P. Johnson, and J. Temple, "Growth Econometrics," University of Wisconsin, Madison, working paper (2004).

Eichen, M., and M. Zhang, "Annex: The 1998 Household Survey: Data Description and Availability," (pp. 331–336), in Keith Griffin and Zhao Renwei (eds.), *Distribution of Income in China* (New York: St. Martin's Press, 1993).

Emran, M. Shahe, and Z. Hou, "Access to Markets ans Household Consumption: Evidence from Rural China" (2008), SSRN Electronic Library.

Emran, M. Shahe, and F. Shilpi, "The Extent of the Market and Stages of Agricultural Diversification," *Canadian Journal of Economics* 45 (2012), 1125–1153.

Emran, M. Shahe, and Yan Sun, "Magical Transition? Intergenerational Educational and Occupational Mobility in Rural China: 1988–2002," paper presented at the AEA Conference (2011).

Frankel, J., and D. Romer, "Does Trade Cause Growth?" *American Economic Review* 89 (1999), 379–399.

Gallup, J. L., J. Sachs, and A. D. Mellinger, "Geography and Economic Development," *International Regional Science Review* 22 (1999), 179–232.

Giles, J., and K. Yoo, "Precautionary Behavior, Migrant Networks, and Household Consumption Decisions: An Empirical Analysis Using Household Panel Data from Rural China," this REVIEW 89 (2007), 534–551.

Hanson, G., "Increasing Returns, Trade and the Regional Structure of Wages," *Economic Journal* 107 (1997), 113–133.

Henderson, J., Z. Shalizi, and A. Venables, "Geography and Development," *Journal of Economic Geography* 1 (2001), 81–105.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

Jacoby, H., "Access to Markets and the Benefits of Rural Roads," *Economic Journal* 110 (2000), 713–737.

Jalan, J., and M. Ravallion, "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China," *Journal of Applied Econometrics* 17(4) (2002), 329–346.

Jian, T., J. Sachs, and A. Warner, "Trends in Regional Inequality in China," NBER working paper 5412 (1996).

Jin, H., Y. Qian, and B. Weingast, "Regional Decentralization and Fiscal Incentives: Federalism, Chinese Style," *Journal of Public Economics* 89 (2005), 1719–1792.

Kanbur, R., and A. Venables, (eds.), *Spatial Inequality and Development* (New York: Oxford University Press, 2005).

Kanbur, R., and X. Zhang, "Fifty Years of Regional Inequality in China: A Journey through Revolution, Reform, and Openness," *Review of Development Economics* 9 (2005), 87–106.

Khan, A. R., and K. Riskin, "Income Distribution in Urban China during the Period of Economic Reform and Globalization," *American Economic Review* 89 (1999), 296–300.

Klein, R., and F. Vella, "Estimating a Class of Triangular Simultaneous Equations Models without Exclusion Restrictions," *Journal of Econometrics*, 154 (2010), 154–164.

Krueger, A., "*Trade and Employment in Developing Countries, Vol. 3: Synthesis and Conclusions*" (Cambridge, MA: NBER, 1983).

Leung, Chi-Keung, *China, Railway Patterns and National Goals* (Chicago: University of Chicago, 1980).

Lewbel, A., "Using Heteroskedasticity to Identify and Estimate Mismeasured and Endogeneous Regressor Models," *Journal of Business and Economic Statistics* 30 (2012), 67–80.

Li, Hongbin, P. Liu, N. Ma, and J. Zhang, "Economic Returns to Communist Party Membership: Evidence from Chinese Twins," *Economic Journal* 117 (2007), 1504–1520.

Little, I., T. Scitovosky, and M. Scott, "*Industry and Trade in Some Developing Countries: A Comparative Study*" (New York: Oxford University Press, 1970).

Mallee, H., "China's Household Registration System under Reform," *Development and Change*, 26 (1995), 1–29.

Mallick, D., "Microfinance and Moneylender Interest Rate: Evidence from Bangladesh," *World Development* 40 (2012), 1181–1189.

Minten, B., and Steven Kyle, "The Effect of Distance and Road Quality on Food Collection, Marketing Margins, and Traders' Wages: Evidence from the Former Zaire," *Journal of Development Economics* 60 (1999), 467–495.

Morduch, J., and T. Sicular, "Politics, Growth, and Inequality in Rural China: Does It Pay to Join the Party?" *Journal of Public Economics* 77 (2000), 331–356.

Murphy, K., A. Schleifer, and R. Vishny, "Industrialization and the Big Push," *Journal of Political Economy* 97 (1989), 1003–1026.

Nee, Victor, "Organizational Dynamics of Market Transition: Hybrid Forms, Property Rights, and Mixed Economy in China," *Administrative Science Quarterly* 37 (1992), 1–27.

Nicita, A., "The Price Effect of Tariff Liberalization: Measuring the Impact on Household Welfare," *Journal of Development Economics* 89 (2009), 19–27.

Ravallion, M., and S. Chen, "China's (Uneven) Progress against Poverty," *Journal of Development Economics* 82 (2007), 1–42.

Redding, S., and Daniel M. Sturm, "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review* 98 (2008), 53–82.

Redding, S., and A. Venables, "Economic Geography and International Inequality," *Journal of International Economics* 62 (2004), 53–82.

Riskin, C., R. Zhao, and S. Li, "*Chinese Household Income Project, 1995*," (Ann Arbor, MI: ICPSP, 2000).

Rozelle, S., "Rural Industrialization and Increasing Inequality: Emerging Patterns in China's Reforming Economy," *Journal of Comparative Economics* 19 (1994), 362.

Rosenstein-Rodan, P. "Problems of Industrialization of Eastern and South-Eastern Europe," *Economic Journal* 53 (1943), 202–211.

Sachs, J., *The End of Poverty: Economic Possibilities of Our Time* (New York: Penguin Group, 2005).

Shea, J., "Instrument Relevance in Multivariate Linear Models: A Simple Measure," this REVIEW 79 (1997), 348–352.

Smith, S., *Ending Global Poverty: A Guide to What Works* (London: Palgrave Macmillan, 2005).

Topolova, P., "Trade Liberalization, Poverty, and Inequality: Evidence from Indian Districts," NBER working paper W11614 (2005).

Whalley, J., and S. Zhang, "A Numerical Simulation Analysis of (Hukou) Labour Mobility Restrictions in China," *Journal of Development Economics* 83 (2007), 391–410.

Winters, L. A., N. McCulloch, and A. McKay, "Trade Liberalization and Poverty: The Evidence So Far," *Journal of Economic Literature* 42 (2004), 72–115.

World Bank, "*Sharing Rising Incomes: Disparities in China*" (Washington DC: World Bank, 1997).

World Development Report, *Attacking Poverty* (New York: Oxford University Press, 2001).

——— *Equity and Development* (New York: Oxford University Press, 2006).

Wu, Yuan-li, *The Spatial Economy of Communist China: A Study on Industrial Location and Transportation* (New York: Praeger, 1967).

Yang, D., "What Has Caused Regional Inequality in China?" *China Economic Review* 13 (2002), 331–334.

APPENDIX

Variables Description

Dependent Variables[a]	
Ci	Household i consumption per capita in 1995 (yuan, deflated by spatial CPI).
Variables of Interest[b]	
Ad	Log of road and railway distance from the county seat of a household to the nearest business center (km)
Ag	Log of road and railway distance from local business center to the nearest major seaport (km)
(Ad) × Ag)	Interaction of distance to the business center and that from business center to the seaport
Controls Used for Results in Table 1[b]	
ter	Type of terrain: mountaineous = 1
rain	Average annual rainfall
temperature	Average annual temperature
elevation	Elevation of the county where household is located (unit: meter)
VAR_EL	Variance of elevation within county
MEAN_SL	Mean slope (in percent, slope = rise/run × 100)
STD_SL	Standard deviation of slopes
RIVER_LG	Total river length (m)
soil	Soil quality
gdp	GDP of household's nearest business center (unit: million yuan)
cargoc	Cargo handled at seaport (unit: ton)
socprod90_d	Average gross output value of society for the intervening counties along the straight-line distance from county seat to navigable river
socprod90_i	Average gross output value of society for the intervening counties along the straight-line distance from business center to coastline

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 1.—(CONTINUED)

Controls at Province Level: Dummies for Nineteen Provinces[b]	
Beijing, Hebei, Shanxi, Liaoning, Jilin, Jiangsu, Zhejiang, Anhui, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu	
Variables in the Z Vector Used in Lewbel (2012) Estimator	
Ddem_1k_std_Wsum	Weighted mean standard deviation of elevations to navigable river (m)
Idem_1k_std_Wsum	Weighted mean standard deviation of elevations to coastline (m)
Idem_1k_mea_Wsum	Weighted mean elevations to coastline (m)
D1	Ratio of the weighted maximum and the weighted minimum of the mean elevation to navigable river
Excluded Instruments for GMM-1 and 2SLS Specification	
Avdist	Straight-line distance from county seat to its nearest navigable river (km)
Idist	Straight–line distance from business center to coastline (km)
AVI	Avdist × Idist
Dch_slop_me_Wsum	Weighted mean slope to navigable river (m)
Ich_slop_me_Wsum	Weighted mean slope to coastline (m)
Dch_slop_su_Wsum	Weighted sum of slopes to navigable river (m)
Ich_slop_su_Wsum	Weighted sum of slopes to coastline (m)
Dch_slop_mi_Wsum	Weighted mean of the minimum slope to navigable river (m)
Ich_slop_mi_Wsum	Weighted mean of the minimum slope to coastline (m)
Dch_slop_su_Wtmax	Weighted maximum of the sum of slopes to navigable river (m)
Ich_slop_su_Wtmax	Weighted maximum of the sum of slopes to coastline (m)
Dch_slop_ra_Wtmin	Weighted minimum of slope range to navigable river (m)
Ich_slop_ra_Wtmin	Weighted minimum of slope range to coastline (m)
Ddem_1k_std_Wsum	Weighted mean standard deviation of elevations to navigable river (m)
Idem_1k_std_Wsum	Weighted mean standard deviation of elevations to coastline (m)
Ddem_1k_max_Wtmax	Weighted maximum elevation to navigable river (m)
Idem_1k_max_Wtmax	weighted maximum elevation to coastline (m)
Ddem_1k_std_Wtmax	weighted maximum standard deviation of elevations to navigable river (m)
Idem_1k_std_Wtmax	Weighted maximum standard deviation of elevations to coastline (m)
delevation	Elevation difference between county seat and business center
ielevation	Elevation difference between business center and seaport
Excluded Instruments for GMM 2 Specification[c]	
Avdist	Straight-line distance from county seat to its nearest navigable river (km)
Idist	Straight-line distance from business center to coastline (km)
AVI	Avdist × Idist
Dch_slop_me_Wsum	Weighted mean slope to navigable river (m)
Ich_slop_me_Wsum	Weighted mean slope to coastline (m)
Dch_slop_ra_Wtmin	Weighted minimum of slope range to navigable river (m)
Ich_slop_ra_Wtmin	Weighted minimum of slope range to coast line (m)
Ddem_1k_std_Wsum	Weighted mean standard deviation of elevations to navigable river (m)
Idem_1k_std_Wsum	Weighted mean standard deviation of elevations to coastline (m)
Ddem_1k_max_Wtmax	weighted maximum elevation to navigable river (m)
Idem_1k_max_Wtmax	Weighted maximum elevation to coastline (m)
Ddem_1k_std_Wtmax	Weighted maximum standard deviation of elevations to navigable river (m)
Idem_1k_std_Wtmax	Weighted maximum standard deviation of elevations to coastline (m)
Delevation	Elevation difference between county seat and business center
ielevation	Elevation difference between business center and seaport
Additional Controls of Individual or Household for Robustness Checks in Table 4	
Sex	Household head's gender
Age	Household head's age
Asq	Age^2
Min	Dummy for ethnic minority status of the household head (minority = 1)
Gov	If household head is village cadre, official of party, or government office institution (yes = 1)
Com	Dummy showing if the household head is a member of the communist party (yes = 1)
Agr	Dummy for household head's occupation (traditional farming = 1)
Edu	Years of schooling of household head
Nchild	Number of children in a household (age $<$ 17)
Nold	Number of old people in a household (age $>$ 60)
Assignedland	Total assigned land used for farming in 1995 (mu)
Initial_assets	Initial fixed productive assets: original value of fixed productive assets at the end of 1994
Initial_livestock	Initial endowment of livestock: original value of livestock at the end of 1994
D_housing	= 1 if house built before 1989, = 0 otherwise
Tap_water	If drinking water obtained from tap (yes = 1)
Lighting	If lighting by electricity (yes = 1)
Bankcredit	Debt borrowed from banks in 1995 (RMB)
Machinery	Original value of industrial machinery and equipment household owned in 1995 (RMB)
Nma	National ethnic minority: yes = 1

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021

TABLE 1.—(CONTINUED)

Additional Controls of Public Goods for Robustness Checks in Table 4	
health_clinic	If there is health clinic in the village (yes = 1)
d_school	Dummy for school in village (yes = 1)
public_F	The employment in public facilities (person, county)
public_H	The employment in public health service (person, county)
public_E	The employment in education (person, county)

[a]Dependent variable described above is used in GMM-1, 2SLS, GMM-2, Lewbel two-stage estimator in table 1, and in level-level functional form in table 4. The log form of dependent variable 1 is used in the log-log functional form in table 4.

[b]Independent variables described above are used in GMM-1, 2SLS, GMM-2, Lewbel two-stage estimator in table 1, and in the log-log and level-level functional forms in table 4 as well. But in the level-level functional form, the independent variables, the variables of interest, are the level values of distances.

[c]The topography features are for the intervening counties along the relevant straight-line distances. They are weighted by using the length of the straight-line distance that passes through the intervening county.

Source for GDP and cargo: China Statistic Yearbook (1996). Sources for road distances, river, elevation contour lines, and other county geography: China Electronic Map 2004; the Digital Chart of China: provided by the CITAS (China in Time and Space–University of Washington) project; and China in Time and Space (1990, 2000), CIESIN, compiled by Robert Dernberger and William Skillers. Sources for county statistics: zhongguo fenxian nongcun jingji tongji gaiyao (1991). The World Bank Group.

Downloaded from http://direct.mit.edu/rest/article-pdf/95/2/682/1917226/rest_a_00354.pdf by guest on 19 September 2021