Political Analysis (2005) 13:1–22
doi:10.1093/pan/mpi001

Reaching Migrants in Survey Research: The Use of the Global Positioning System to Reduce Coverage Bias in China

Pierre F. Landry
Yale University, Department of Political Science,
124 Prospect Street, New Haven, CT 06520-8301
e-mail: pierre.landry@yale.edu

Mingming Shen
Department of Political Science, School of Government,
416 Library Building, Peking University, Beijing, China 100871
e-mail: shenmm@rcccpku.org

List-based samples are often biased because of coverage errors. The problem is especially acute in societies where the level of internal migration is high and where record keeping on the population is not reliable. We propose a solution based on spatial sampling that overcomes the inability to reach migrants in traditional area samples based on household lists. A comparison between a traditional study and our sample of Beijing demonstrates that coverage bias is greatly reduced. The successful incorporation of mobile urban residents has important substantive effects, in both univariate and multivariate analyses of public opinion data.

1 Introduction

The use of accurate population lists is a sine qua non condition for the design and use of probability samples in the social sciences. For general social surveys, lists are typically based on publicly available census data or directories. A census, however well executed it may be, almost always produces undercounts (USCMB 2000). Directories are also rarely complete enumerations of the populations to be surveyed (Kviz 1984). Since excluded populations tend not to be missing at random, the resulting census or survey data is usually biased, particularly with respect to minorities, immigrant groups, or the homeless (Subcommittee on Census and Population 1987; Skerry 2000; Anderson and Fienberg 2001; Steenkamp and Van Aardt 2001).

Authors' note: This research is a collaborative endeavor between Yale University and the Research Center for the Study of Contemporary China at Peking University. Financial support by the Yale Center for International and Area Studies and Social Science Research Grant at Yale is deeply appreciated. Our gratitude goes to the staff of the RCCC for their hard work, dedication, and creativity. Thanks to Chris Achen, Chai Jingjing, Deborah Davis, Steve Harringer, Hao Hongsheng, Tom Piazza, Kenneth Scheve, Marty White, Yan Jie, and Yang Ming and for their invaluable suggestions at various stages of this project. We are also grateful for the comments received on an earlier version of this article presented at MPSA in April 2003 and the Yale faculty workshop, as well as those of the anonymous reviewers. All errors remain ours.

The problem of incomplete coverage has a long pedigree in survey research. In developed countries, survey biases induced by inaccurate lists can be compensated using random digit dialing (RDD) techniques to reach respondents by telephone. The validity of this approach depends crucially on two factors: the extent of phone penetration among households and the avoidance of nonresponse bias once respondents have been reached (Tull and Albaum 1977). However, in combination with lists, RDD techniques can be reliable (Brick et al. 1995).

These techniques are not appropriate in developing societies because phone penetration is still low and remains heavily biased toward affluent, typically urban, populations. Furthermore, in societies experiencing high levels of internal population movement, migrants are far less likely to have a telephone than the rest of the population. In such cases, official lists, however imperfect, are still more reliable that RDD designs but still fail to produce unbiased samples.

In this study, we demonstrate how coverage bias can be reduced with respect to a specific situation, internal migrants in China. Yet similar problems exist in many contexts: one may want to reach undocumented immigrants in metropolitan areas of developed countries, survey refugee camps, or war-torn countries for the purpose of planning emergency assistance, or simply draw samples in countries that have not conducted a census in decades, like Iraq or Afghanistan. Our experiment was undertaken in an area where much official data was available and was used in order to validate our approach to global positioning system (GPS) sampling. However, the methodology can be easily extended to a wide variety of situations in which researchers require equal probability samples but are unable to access reliable lists of the populations of interest.

2 Survey Research in China: Accounting for 140 Million Migrants

In China, the collection of census data depends largely on the cooperation of institutions in charge of population control: since 1958, all urban residents have been expected to register their presence with local neighborhood committees (NCs)[1] and obtain a formal household registration.[2]The combination of neighborhood committees and household registrations proved to be a highly effective barrier to entry into cities. Unlike many developing countries, China did not experience mass migration from rural areas to urban centers during its initial phase of industrialization. Instead, Chinese society was artificially divided into two theoretically distinct spheres. Outsiders could not decide to settle in urban areas without obtaining a proper *hukou* granting them access to a slew of benefits of the urban regime, including grain. Only under exceptional circumstances could citizens transfer their registrations from rural to urban areas (Goldstein 1987; Gui and Liu 1992; Johnson 1994; Solinger 1995, 1999; Goodkind and West 2002; Wu and Treiman 2004).

Because the NCs were fairly efficient collectors of population data, household registration lists have been widely used as the basis of the vast majority—if not all—of probability social samples taken in China since 1978. However, China's rapid economic transformation is affecting the quality of survey research. Until the early 1990s, reasonable lists of residents could still be obtained from neighborhood committees, but market reforms have severely undermined the oversight capacity of NCs. With the end of the rationing of grain and various necessities in urban areas, rural residents, who would not have survived in a city without proper registration under the planned economy, began to

[1]These are called *jumin weiyuanhui* in Chinese. Villages perform similar functions in rural areas.
[2]Or *hukou.*

Table 1 Internal migration in China, by province (year 2000)

	Total Population (million)	*Migrants (%)*	*Place of Origin*		
			Own county	*Own province*	*Out of province*
Beijing	13.6	34%	15%	1%	18%
Shanghai	16.4	33%	6%	7%	19%
Guangdong	85.2	30%	5%	7%	18%
Tianjin	9.8	22%	14%	1%	7%
Zhejiang	45.9	19%	7%	4%	8%
Fujian	34.1	17%	6%	5%	6%
Neimenggu	23.3	16%	9%	5%	2%
Liaoning	41.8	15%	10%	3%	2%
Xinjiang	18.5	15%	5%	3%	8%
Hainan	7.6	13%	4%	4%	5%
Jiangsu	73.0	12%	6%	3%	3%
Ningxia	5.5	12%	6%	3%	3%
Shanxi	32.5	11%	7%	2%	2%
Jilin	26.8	11%	7%	2%	1%
Qinghai	4.8	11%	4%	4%	3%
Heilongjiang	36.2	10%	5%	4%	1%
Hubei	59.5	10%	6%	3%	1%
Yunnan	42.4	9%	3%	3%	3%
Chongqing	30.5	9%	6%	2%	1%
Jiangxi	40.4	8%	6%	2%	1%
Shandong	90.0	8%	5%	2%	1%
Tibet	2.6	8%	2%	2%	4%
Sichuan	82.3	8%	5%	3%	1%
Guangxi	43.9	7%	3%	3%	1%
Hebei	66.7	7%	4%	2%	1%
Hunan	63.3	7%	4%	2%	1%
Guizhou	35.2	7%	3%	2%	1%
Shaanxi	35.4	7%	4%	2%	1%
Gansu	25.1	6%	3%	2%	1%
Anhui	59.0	6%	4%	2%	0%
Henan	91.2	6%	3%	2%	1%
CHINA	1242.6	12%	5%	3%	3%

Source. Computed from SSB (2002).

migrate to urban centers in search of new employment opportunities. Individual entrepreneurs, collectives, and even state firms began to offer temporary employment to rural residents who were willing to take the risk of violating registration regulations. Employment agencies now routinely bring workers to cities, sometimes from very remove parts of the country.

The erosion of the household registration system has had massive consequences. The latest census, conducted in 2000, classified 144.4 million people as migrants (11.62% of the total population). Sixty-five million migrated within their county of origin, 36 million left their home county but settled in their home province, and 42 million crossed provincial boundaries (SSB 2002, Tables 1 and 5). In the coastal regions of Beijing, Tianjin, Guangdong, and Shanghai, the proportion of migrants currently exceeds 20% of the population.

The emergence of a large migrant population has been widely noted (Goldstein 1987; Yang 1993; Solinger 1995, 1999; Zhang 2001) but has not resulted in serious reevaluation of survey research based on sample frames derived from NC lists. A number of studies have attempted to reach different types of migrants. Goldstein et al. (1991) were able to analyze migrants who reside in formally registered households, but they concede that this approach restricts the analysis to a very specific subset of migrants. As Goodkind and West (2002) discuss, migrants include a vast array of types, ranging from long-term residents who were ultimately successful in obtaining a formal registration to the "floating population" of informal, short(er)-term migrants who may return to their home towns quite regularly.

Short-term migrants have been systematically excluded from the few studies that have sought to quantify the extent of migration. In 1987, the 1% Population Survey of China defined eligible respondents as persons who had lived in the place of interview for at least six months. Short-term residents were not considered. Similarly, the 1990 population census used a one-year benchmark (Liang and White 1996, p. 376 n.3). Liang and White (1996) base their estimates on a 10% random sample of the 2% Fertility and Birth Control Survey conducted by the State Family Planning Commission in 1998. That survey collected information on past interprovincial migration decisions among all members of the households surveyed, but the study does not account for intra-provincial migration. The 1995 1% population survey had a more sensible design: all household types were accounted for because the design called for a 1% clustered sample of village-level units: once a village was selected, all households in the village were surveyed, regardless of registration status. The tabulation can thus separate responses by registration category (SSB 1997; Fan 2000, n.10).

These studies all attempted to estimate the magnitude and spatial characteristics of migration. The best designs (such at the 1995 1% sample) provide reasonable measures of the "problem" but provide no information about the characteristics, attitudes, and behavior of migrants on many variables of interest to social and political scientists. Only aggregate tabular data is publicly available, precluding even simple cross-tabulations of the results.

These studies are also suspiciously dependent on official data and government institutions. Most surveys (even the census) ultimately depend on the collaboration of neighborhood committees that are supposed to provide demographic information to samplers and enumerators. The results are predictably suspicious: Fan, for example, reports, implausibly, that migrants have much higher education than nonmigrants (Fan 2000; Table 1). Similarly, SSB's 1995 1% sample puts the total migrant population at 35.3 million, far below conventional projections (Liang and White 1996).

Thus survey researchers who rely on registration lists face a serious inferential threat: because up to a quarter of China's de facto urban residents are *not* formally registered, probability samples based on NC lists are only probability samples of a *subset* of urban residents. They no longer cover the entire population of their sample space.

3 Reaching Migrants

The methodology presented in this article was used to address the specific problem of unaccounted migrants in urban China, but it could be an effective tool to administer social surveys in an array of challenging research environments. Survey research in many developing counties is hampered by the lack of complete and accurate population data. In the absence of census data, equal probability samples of Kabul, Phnom Penh, or Kinshasa are impossible unless researchers first enumerate these spaces themselves, which is

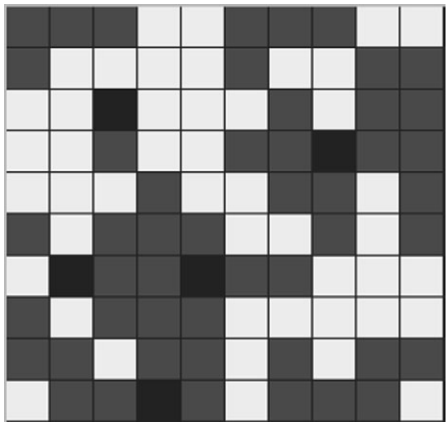


Fig. 1 Hypothetical spatial sample frame
Unit types: U_1 (black) U_2 (dark grey) U_3 (white)
$T = \sum \phi_i t_i = 100$
$\phi_1 = 5/100$
$\phi_2 = 48/100$
$\phi_3 = 47/100$

obviously impractical. The approach taken here is a radical departure from the traditional methodology of area sampling: rather than assuming that the population is known ex ante, we derive a spatial sampling protocol that still produces equal probability samples of households as long as a small number of operational principles are properly implemented. Information about the demographic characteristics of the population is useful—but not necessary—insofar as it helps determine the magnitude and cost of fieldwork and forecast sample size. Properly administered, spatial samples can provide a *universal* tool to produce equal probability samples of any population, regardless of the quality of the census data or any "official" population count.

3.1 *The Basics*

Consider a hypothetical physical sample space in which a population is distributed over T units, such as the stylized space in Fig. 1. Assume (for now) that we can obtain ex ante measures of size for all unit types. Let ϕ_i represent the fraction of units with population density i. ($\sum \phi_i = 1$)

3.1.1 Equal Probability Sampling with ex ante Measures of Unit Size

Like any sample, a random sample of spatial units produces an unbiased distribution of the underlying population: if the sample is large enough, the expected proportion of each unit type is equal to ϕ_i, and sampling error is a function of the sampling ratio. If we assume that each spatial unit is populated with households, we can compute the probability of selecting a given household, conditional on the probability of selecting a spatial unit.

Let k be the number of spatial units that we wish to include is our sample. The probability of selecting a unit is $\pi = k/T$. Let θ be the probability of selecting a household within a unit ($\theta = n_j/N_j$ where N_j is the number of households that reside in that unit).The probability of selecting a household is simply the product of π and θ:

$$\Pi = \pi \cdot \theta = \frac{k \cdot n_j}{T \cdot N_j}.$$

If we wish to account for uneven densities across spatial units, the probability of selecting a household in a spatial unit of type i is simply $\phi_{i\cdot} \cdot \Pi$.

$$\Pi_{ij} = \phi_i \pi\theta = \phi_i \frac{k \cdot n_j}{T \cdot N_j}.$$

Thus in the hypothetical case of a 10×10 frame with five densely populated areas, if $k = 10$, the probability of selecting one household in a densely populated area (black) is (5/100)*(10/100)*(1/N) = .005/N

If *k>1,* selected units may have uneven densities, thus the probabilities of selecting households vary with unit size. When reliable population data is available, we can recover equal probabilities by choosing unit-specific n_j or use sampling weights during the analysis.

3.1.2 Equal Probability Sampling without ex ante Measures of Unit Size

If N is unknown ex ante, having drawn a unit with probability π, we can ensure equal probability of selection among households by setting $n{=}N$: each household within a given sampling unit is interviewed with certainty:[3]

$$\forall N_j, \Pi = \pi\theta = \frac{k \cdot N_j}{T \cdot N_j} = \Pi.$$

This result is important for two reasons. First, it is *independent of population measures.* The probabilities of selection depend solely on the proportion of spatial units that are included in the sample. Second, the probabilities are *invariant with the shape or surface* of the spatial units. All that is required is a clear definition of unit boundaries (on a map) so that, if selected, the research team can interview all respondents who live in that space, or a fixed proportion thereof.

The requirements are very similar to block listing: ultimately, once an area is randomly selected, it must be enumerated. The cost and data requirements of traditional block listing are very high, particularly in countries where "blocks" have no statistical definition. Clear delineation of natural blocks can be very challenging: in cities like Beijing, old courtyards in the traditional Beijing neighborhood structures, known as *hutong* in Chinese, are so imbricated that reasonable blocks can easily include dozens of households. The process is also time consuming and expensive. With limited budgets, the research team may be able to afford block listing in only a small number of neighborhoods; this adversely affects the geographical representativeness of the sample.

3.2 *Feasible Block Listing with GPS Technology*

The spatial approach is more efficient than block listing. Rather than enumerating unwieldy blocks, one can grid the sample space in such a way that each unit is small

[3]If the cost of interviewing all households is too high, an alternative is to draw spatial units and then select a fixed proportion of households across units. Such an approach has the advantage of reducing clustering effects among respondents and freeing up resources that can to used to increase the number of spatial sampling units, which improves the geographical representativeness of the sample

enough to be enumerated quickly and cheaply. For a given expected sample size, the risk of producing unrepresentative samples is reduced: smaller units imply that one can incorporate more neighborhoods.

In studies that rely on block listing, block size is largely determined by the convenience of locating them. Interviewers can be told to enumerate a list of addresses inside a well-defined geographical area bound by streets (with names) or some other easily identifiable demarcation. Arbitrarily defining small blocks independent of natural boundaries was until recently very difficult because such units were very hard to reach in practice. For instance, one could not typically instruct an interviewer to visit a site centered on "North 39 degrees, 10 minutes and 25.76 seconds, East 111 degrees, 24 minutes and 45.12 seconds" unless she happened to be trained as a professional land surveyor. The technical expertise, costs, and time needed to accomplish this task were simply prohibitive.

Modern GPS receivers solve this problem, since they have the ability to identify small units with considerable precision almost anywhere on the planet. We can, for instance, select a "block" defined as a square second of area at random and use a GPS receiver to guide us to our destination with a precision of a few meters. Surveyors may then enumerate the households residing within the boundaries of the square second. Once listed, each household (or a fixed proportion of households to respondents) must be interviewed.[4]

Alas, there is a practical drawback: since the probabilities of selection depend not on the number of households but on the number of spatial points (k/T), the *final sample size cannot be determined ex ante*. We can only rely on "best estimates" of population densities to determine an appropriate number and size of final sampling units.

3.3 *Sampling Design, in Detail*

As a comparative benchmark, we use a traditional equal probability sample of Beijing (registered) households, as listed by neighborhood committees, which was conducted by the RCCC in December 2001.[5] In order to fairly compare our findings with the results of the traditional study, we designed a spatial grid of Beijing to match the geographical expanse of the Beijing Area Study (BAS). [6] Populated points that lie outside the BAS frame are considered empty, but all residents of a street committee listed in the BAS sample frame are eligible, regardless of their registration status.[7]

Given the need to control fieldwork cost and time, we also used the grid to create a rough population model of Beijing. Grid cells represent spatial "square minutes." Each cell was coded to determine the relative contribution of relevant *jiedao* to the surface of the primary sampling unit (PSU), defined as the surface of constructed zones plotted on the latest available street map of Beijing (Fig. 2).

3.3.1 Correcting for Migrants

Since our initial codings used the BAS sample frame to estimate the registered population, they had to be further adjusted to account for migrants in each PSU. At the time, this

[4]We used standard Garmin E-Trex receivers, which cost about $120 each.

[5]The Beijing Area Study (BAS) is an annual survey of the eight urban districts of Beijing that has been conducted since 1995 by the Research Center for Contemporary China at Peking University (RCCC).

[6]In China, neighborhood committees are under the jurisdiction of street committees (*jiedao banshichu*). Some Beijing districts such as Haidian and Chaoyang include a mix of urban committees (*jiedao*) and townships (*zhen*). Since the latter usually denote peri-urban "rural" areas, they are excluded from the sample frame.

[7]In practice, an eligible respondent was defined as an adult 18 to 65 years old who had been in Beijing for at least one month on the day he was interviewed. Thus tourists, businessmen, and short-term visitors who do not actually live in the city were excluded.

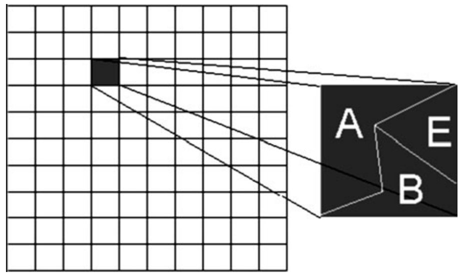


Fig. 2 Coding Rules for PSU: An ExampleE= Empty Space (about 20% of the surface of the square)A= Street Committee A (about 50% of the surface of the square)B= Street Committee B (about 30% of the surface of the square)At Beijing's latitude, a "square minute" is in fact a rectangle measuring 2.62 km^2. Thus, given the population density of each SC, we can estimate the population of the PSU.

information was not available at the level of the street committee.[8] We relied instead on estimates of annual statistics published by the Beijing municipal government, aggregated by urban district.

Each PSU estimate was then multiplied by a correction coefficient that accounts for the likely proportion of migrants in the district where the PSU is located. Thus, if a PSU located in Dongcheng was initially coded with a registered population of 10,000, its total population estimate is $10{,}000 \times 1.15 = 11500$ individuals. We therefore made the (strong) assumption that the density of migrants varies across districts but that migrants are uniformly distributed in space within each district. Because the published data were aggregated at the level of the urban district, we also introduced some aggregation bias, since the density of migrants is lower in the rural townships of these districts that are excluded from the sample frame.

One might wonder whether our reliance on government data to estimate densities and measures of size is warranted, since it is precisely the flaws of official data that motivated our decision to experiment with GPS sampling in the first place. In defense of Chinese statistics, it is important to note that the census did produce estimates of the migrant population. In our view, the data are flawed because only long-term migrants were counted. Yet when computing measures of size, we used the 2000 census migration data in order to calculate our final estimates of the total population. Recall that in PPS sampling, only *relative* MOS matter. Thus our correction coefficient introduced some bias to the extent that the ratio of short-term to long-term migrants varies across units.

This is why we underestimated the proportion of migrants within the sample space. Table 2 summarizes one measure of the magnitude of the error by comparing the data used at the time of the study (column A) with the correction coefficient that would have been used has the census data been available at the time (Column B). Migrant density is higher in all districts, but the errors are most pronounced in the districts that form the immediate outer ring of the city (Chaoyang, Fengtai, Haidian, and Shijingshan), while the errors in the central business district are considerably lower.[9]

[8]The proportion of migrants by street committee/township was not available at the time but has since been published (see SSB 2002).

[9]The discrepancy between column A and column B is not necessarily a case of statistical manipulation. Column A includes the rural townships of each urban district, while column B restricts the computation to the street committees that lie within the BAS sample frame and therefore excludes townships.

Table 2 Correction coefficient to account for migrants, by urban district

	A	*B*	*B-min*	*B-max*
Dongcheng	1.15	1.40	1.34	1.50
Xicheng	1.10	1.50	1.39	1.57
Chongwen	1.13	1.32	1.28	1.52
Xuanwu	1.15	1.33	1.19	1.86
Chaoyang	1.29	1.31	1.27	2.97
Fengtai	1.39	2.64	1.20	7.97
Shijingshan	1.27	1.66	1.31	3.82
Haidian	1.25	1.20	1.26	5.30
BEIJING	1.24	1.40	1.19	7.97

A=Based on Beijing Statistical Yearbook 2001
B=Based on 2000 census (street committees only)

Since census data have now been released at the township level, we can measure the impact of the forcing migrant density in all areas on the mean value in the urban district. Again, this assumption was acceptable in the city center but can lead to serious forecasting errors in neighborhoods where migrants are heavily concentrated. In one *jiedao* in Fengtai, the ratio of migrants to registered households is 8 to 1 (Fig. 3).

The final grid consists of 349 PSUs (square minutes). At Beijing's latitude, a square minute is in fact a quasi-rectangular trapezoid 1.852 km tall (latitude) and 1.418 km wide (longitude), namely a 2.262 km^2 neighborhood.[10]

Figure 4 summarizes measures of size estimates. Expected population is plotted along the vertical axis, arranged by latitude and longitude. Note that we did not produce a model of the population of Beijing as a whole; we sought only to estimate population density in the area of the city covered by the baseline study.

We treated each square minute as a PSU for first-stage selection of neighborhoods and used our measures of size to sample 50 units by PPS (probabilities of selection proportional to measures of size). In order to ensure adequate representation of urban as well as peri-urban neighborhoods, the PSUs were sorted in decreasing order of measures of size. Thus all households were selected with equal probability, but the design guaranteed the selection of a representative mix of low-, medium-, and high-density neighborhoods.

3.3.2 Secondary Sampling Units: Square Seconds

Next, we selected 200 secondary sampling units (SSUs) at random defined as "square seconds" (four SSUs drawn completely at random within each PSU). In order to facilitate

[10]This estimation is based on the formula for the computation of great arc distances at sea level, assuming a spherical Earth.
The distance between two points (A,B) is:

$dLONG$ = longitude B − longitude A

$dLAT$ = latitude B − latitudeA

$a = \sin(dLAT/2)^2 + (\cos(\text{lat}1) \cdot \cos(\text{lat}2) \cdot \sin(dLAT/2)^2)$

$c = 2 \cdot \arcsin(\min[1, \sqrt{a}])$

d = R· c where R is the average between the polar radius and the equatorial radius of the earth (R ≅ 3956 miles ≅ 3365 km)

(See Chamberlain 2002.)

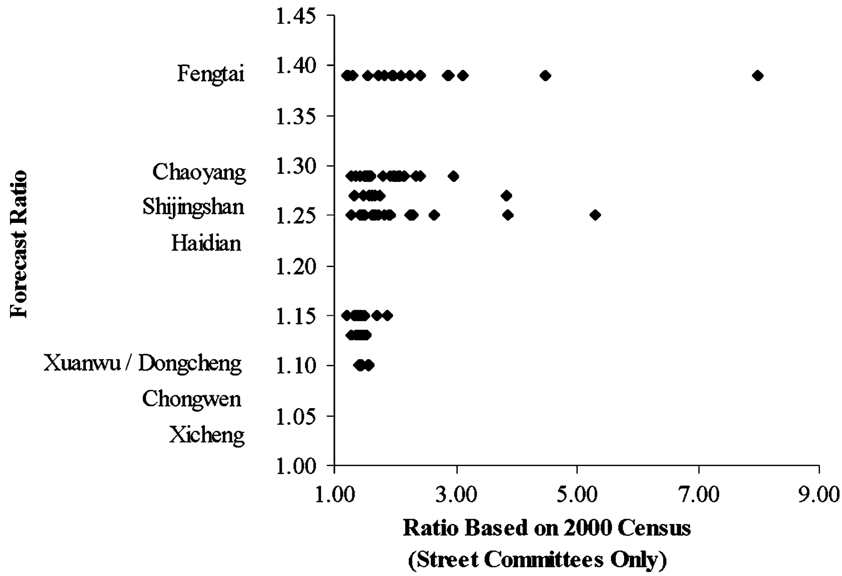


Fig. 3 Correction coefficient to account for migrants: comparison between municipal and census data. Note: Each dot represents as street committee (horizontal axis). Forecasts vary across, but are constant within, urban districts (vertical axis).

the task of surveying SSUs, the teams did not locate elongated trapezoidal square seconds. Instead, they mapped actual squares of identical surface (54×54 meters), centered on the geo-coordinates of the related square second and enumerated all households in these units.

Having surveyed all 200 points,[11] we realized that we had counted many more "doors" than expected, four times what was needed to reach 1000 households. This error was most likely due to the highly heterogeneous densities of the urban space. In Beijing, half of the square minutes surveyed were completely empty, while the nonempty ones included many more doors than our measures of household density implied. Furthermore, we were comparing doors to expected households, assuming a one-to-one relationship, which is only a very rough approximation of the housing reality of Beijing residents.

We also based our expectations of sample size on the best maps and statistical data available, but these tools seem appropriate for rough aggregate estimates. We are reasonably confident that our relative measures of PSU size were acceptable, but they are insufficient to make precise predictions of sample size at the level of the square second. The uncertainty of relationship between doors and households as well as the difficulty of correctly forecasting the number of migrants both contributed to this forecasting error.

Since the number of doors was about four times larger than our budget permitted, we selected one-fourth of the square seconds at random, namely a single point per PSU. Having verified that 24 of these were empty (based on surveying work), we returned to 25 nonempty squares to administer the questionnaire. This decision was consistent with the principle of maintaining equal probability of selection across households, but it illustrates

[11]One PSU was later excluded from the dataset because we determined that all corresponding SSUs were outside the BAS sample frame. Thus we used 49 PSUs.

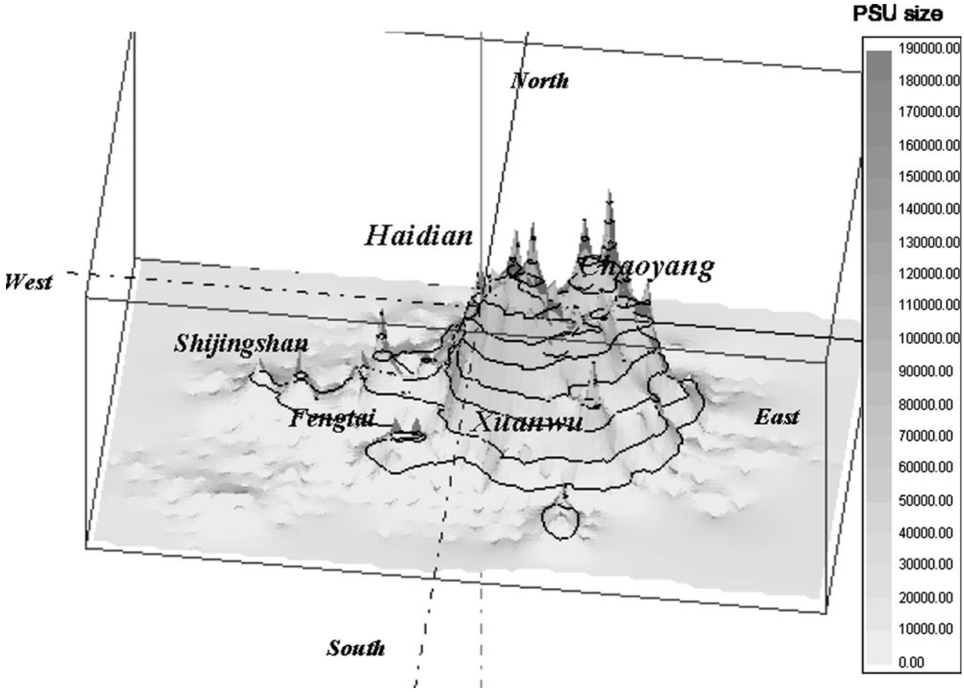
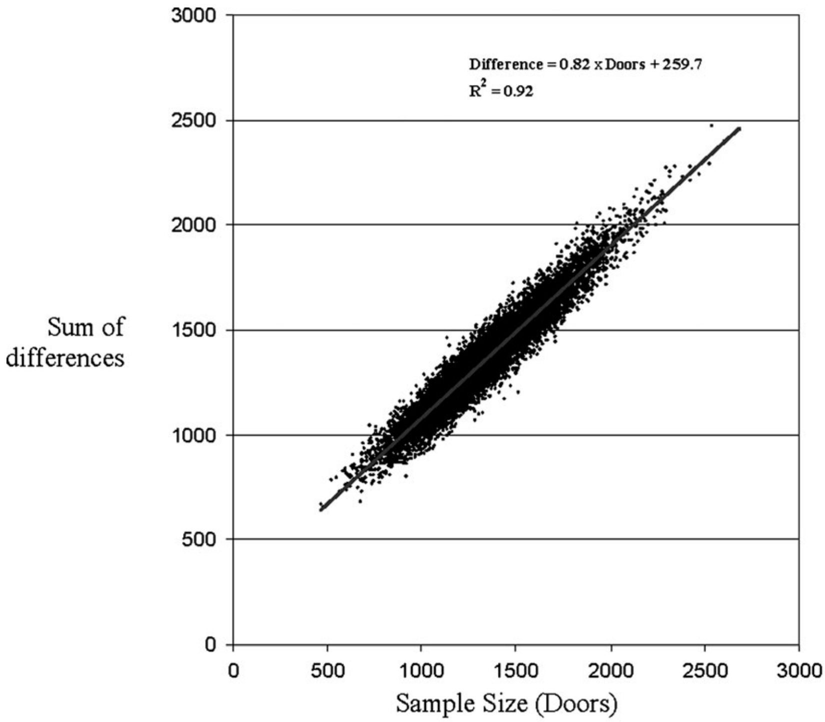


Fig. 4 Population Model of the BAS sample frame Note: Peaks denote the estimated population of each eligible square minute, plotted vertically

the difficulties of this kind of fieldwork. The phases of surveying sampling areas and interviewing respondents had to be distinct: we had to wait until all areas had been surveyed to count the total number of doors and compute the feasible sampling ratio of squares (1:4).

Since we ultimately used only a subset of 200 points surveyed, we can simulate the variance of sample size as well as the extent of the gap between the expected number of households and the numbers of doors that were actually counted. The simulation assumes that we randomly draw one point in each PSU, with 10,000 replications. For each draw, we compute the sample size (total number of doors) and the sum of the differences between expected number of households (from the population model of Beijing) and the actual number of doors across PSUs (Fig. 5).

By construction, the average sample size of 1348 is our target sample size. Here we are instead concerned with sample variance. Given the uncertain relationship between households and actual doors, it is important to measure the likelihood of drawing a sample large enough that meaningful analysis can be conducted, but not so large that we need to—as turned out to be the case—resample SSUs and "waste" surveying time and expenses. There is a 93.4% chance of drawing a sample between 800 and 1800 respondents, 69.1% between 1000 and 1800, and 60.1% between 1000 and 1500. We believe that such odds are reasonable to plan fieldwork. Furthermore, with experience, the model used to compute household estimates can be refined and the number of sites needed to reach our target sample size rescaled in light of these findings. With 49 points, we ultimately completed 738 interviews, slightly more than BAS. We are confident that 50 points constitute an appropriate baseline for future studies in Beijing.



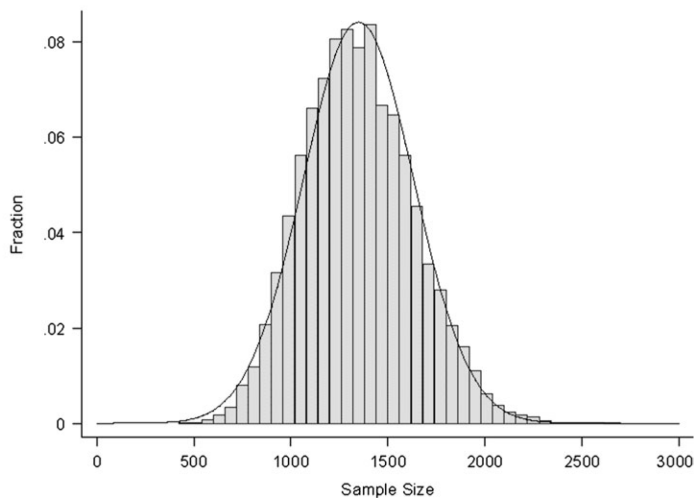


Fig. 5 Simulated Sample Size vs. Difference between Expected Number of Households and Doors actually counted. Note: 10000 draws of ¼ of the points in 49 PSUs actually surveyed for the study Many thanks to Robert S. Kissel for writing the simulation program.

4 Does GPS Sampling Work?

The success or failure of the experiment depends on three factors: (1) the ability to cover the population better than existing alternatives, (2) the replication of basic results for the subset of populations surveyed in this experiment as well as previous BAS studies, and

Table 3 Composition of the BAS-GPS sample, by registration status

	Number	*Percentage*
Migrants	189	25.61
Movers	146	19.78
Correct registration	403	54.61
Total	738	100.00

finally (3) the sensitivity of sampling error to the design of the GPS sample relative to the traditional (simpler) design.

4.1 *Coverage*

The GPS sample undoubtedly has vastly different characteristics than the design based on household registration lists. A stunning proportion of the 738 completed interviews (about 45%) could not have been reached using the traditional method: 25% are migrants who were previously excluded a priori. In addition, 20% of the respondents are "movers," namely formal Beijing residents who do not actually live in the neighborhood in which they are registered (Table 3). In BAS, such movers are almost always treated as "survey nonresponse": if a mover is randomly selected, she could not be reached in practice.

Furthermore, the proportion of migrants is somewhat larger than municipal government estimates imply: assuming the overall correction coefficient of 1.25, for 547 respondents (146 + 403) who hold a Beijing *hukou*, we would be expected to reach about 137 migrants [i.e., $(547 \times 1.25) - 547$]. Instead, the GPS sample includes 189 migrants, 52 more than expected. Our results are therefore consistent with the discrepancy between the municipal data that we used to forecast sample size and the census data. Since published estimates tend to underestimate the true extent of migration into Beijing, we ought to draw more migrants than predicted, and we did. Figure 6 illustrates the extent of these differences for all PSUs, similarly computed with district-level correction coefficients. Simple bivariate results confirm that the official undercount of migrants is systematic. This finding is hardly surprising: local officials have few incentives to report the magnitude of a phenomenon that they are expected—but increasingly unable—to control. Conversely, migrants who face the possibility of sanctions (if not outright expulsion from the city in extreme cases) have no incentive to reveal their presence to authorities. Nevertheless, the overall composition of the sample strongly suggests that we have successfully covered all types of residents, regardless of their registration status. Figure 6 is also reassuring because relative measures of size were generally preserved. The absolute forecasts may have been incorrect, but neighborhoods where we expected to reach fewer migrants than elsewhere are indeed areas where observed densities were low.

4.2 *Comparison with the BAS Sample*

Beyond coverage, the quality of the results also depends on the degree to which the Beijing GPS can replicate BAS findings for groups that are common to both designs, namely residents who actually live in the neighborhood where they are registered.

4.2.1 Accounting for Design Effects

Valid comparisons across samples must account for sampling design (Kish 1965, 1995). Stratification is not an issue in this case: the BAS design simply calls for PPS selection of all neighborhoods in the urbanized section of Beijing's eight districts. BAS-GPS follows

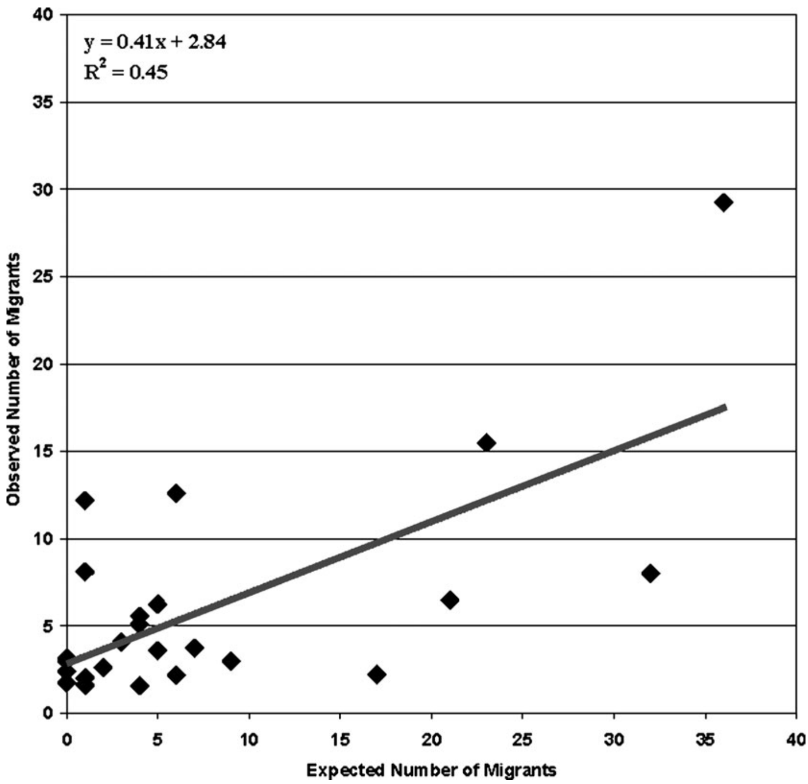


Fig. 6 Relationship between Expected and Observed Number of Migrants, by PSU (BAS-GPS)

Regression:

F(1, 20) = 40.11 Prob > F = 0.000

R^2=0.4556 Root MSE = 7.6433

Dependent Variable: Observed Number of Migrants N=25

	Coef.	Robust SE
Expected Number of Migrants	1.12	0.18 ***
Constant	0.85	1.48

the same approach: all PSUs are drawn from a single stratum that matches the geographical expanse of previous studies.

Probability weights are not required in BAS because the lists supplied by neighborhood committees allow drawing respondents directly, thus avoiding the step of selecting a household, and then randomly choosing an eligible respondent. The GPS method is more complex. All addresses in SSUs (square seconds) must be interviewed to ensure equal probability of selection across households, but only one eligible adult per household can be a respondent. Individual weights must therefore account for variations in household size as well as the number of households that each interview "represents."[12]

[12]Specifically, w=k/p, where $p = \frac{\hat{N}_{iPSU}}{\sum_1^{349} \hat{N}_{iPSU}} \cdot \frac{50}{349} \cdot \frac{1}{3600} \cdot \frac{1}{N_{household}}$.

$N_{household}$ is the number of eligible adults in the respondent's household, 1/3600 is the probability of selecting a square second in a square minute, 50/349 is the proportion of sampled PSUs, and $\hat{N}_{iPSU} / \sum_1^{349} \hat{N}_{iPSU}$ is the share of the PSU (square minute) in the expected population of the sample frame. *K* rescales the weight to compensate for the disparity between expected and observed households.

Table 4 Cross-sample comparisons of respondents' age (BAS vs. BAS-GPS)

Study	*Estimation Procedure*	*Mean*	*Standard Error*	*95% Conf. Interval*		*Deff*	*N*	*PSUs*
AGE								
BAS	Corrected for Design Effect	43.1	0.47	42.1	44.0	0.96	615	46
BAS	Uncorrected (completed interviews)	43.1	0.48	42.1	44.0		615	
BAS	Uncorrected (all sampled)	42.5	0.36	41.8	43.2		1045	
-All Respondents Combined								
BAS-GPS	Corrected for Design Effect	38.4	1.53	35.2	41.5	9.70	731	25
BAS-GPS	Uncorrected	39.2	0.49	38.3	40.2		731	
-Migrants								
BAS-GPS	Corrected for Design Effect	29.4	1.28	26.8	32.1	2.74	189	21
BAS-GPS	Uncorrected	30.1	0.78	28.6	31.7		189	
-Correct Household Registration								
BAS-GPS	Corrected for Design Effect	43.9	1.00	41.9	46.0	2.62	402	23
BAS-GPS	Uncorrected	43.7	0.62	42.5	44.9		402	
-Movers								
BAS-GPS	Corrected for Design Effect	37.8	1.24	35.2	40.4	1.66	146	20
BAS-GPS	Uncorrected	38.8	0.97	36.8	40.7		146	

Clustering occurs in both designs: by neighborhood committee for BAS and by square second with GPS sampling. Therefore, unless otherwise specified, computation of design effects for BAS merely account for clustering by PSU, while design effects for GPS incorporate clustering by PSU and probability weights.

4.2.2 Testing Design Effects Demographic, Economic and Political Variable

Demographics. Since the fieldwork for these studies was conducted in two phases seven months apart, it is preferable to first compare results and design effects for variables that are not time sensitive. Due to space limitation, we discuss only gender and age.

Table 4 illustrates the dramatic impact of GPS sampling on the demographic profile of the sample. Accounting for design effects, the BAS average respondent was 43.1 years old in December 2001. In the GPS study, the mean age is merely 38.4. All of the difference can be explained by the better coverage of the spatial design. The mean age is identical between BAS respondents and comparable to the subpopulation of the BAS-GPS (respondents who hold a Beijing *hukou* and live in the neighborhood where they are registered). In contrast, the GPS sample accounts for the much younger migrants (typically in their late 20s) as well as within-city movers who are six years younger than other *hukou* holders.

Design effect *(Deff)* measures a point at the partial trade-off between the two approaches: variance estimates are lower in the traditional design, since probability weights are unnecessary. However, the ability of spatial sampling to properly predict mean values for respondent types covered in both designs and to account for the vast differences in mean values among groups that are otherwise not covered (migrants) or unreachable (movers) strongly favors spatial sampling. There is little point in generating efficient but biased estimates.

Similar conclusions apply to gender: sampling respondents from *hukou* lists tends to yield too many males, whereas 52% of the BAS-GPS respondents are women. However, at

Table 5 Cross-sample comparisons of respondents' gender (BAS vs. BAS-GPS)

Study	*Estimation Procedure*	*Mean*	*Standard Error*	*95% Conf. Interval*		*Deff*	*N*	*PSUs*
GENDER (1=Female)								
BAS	Corrected for Design Effect	0.46	0.02	0.42	0.50	0.97	615	46
BAS	Uncorrected (completed interviews)	0.46	0.02	0.42	0.50		615	
BAS	Uncorrected (all sampled)	0.47	0.02	0.44	0.50		1072	
-All Respondents Combined								
BAS-GPS	Corrected for Design Effect	0.52	0.04	0.44	0.59	3.62	738	25
BAS-GPS	Uncorrected	0.51	0.02	0.47	0.55		738	
-Migrants								
BAS-GPS	Corrected for Design Effect	0.53	0.02	0.49	0.56	0.63	549	23
BAS-GPS	Uncorrected	0.51	0.02	0.47	0.56		549	
-Correct Household Registration								
BAS-GPS	Corrected for Design Effect	0.53	0.02	0.48	0.58	0.97	403	23
BAS-GPS	Uncorrected	0.52	0.02	0.47	0.57		403	
-Movers								
BAS-GPS	Corrected for Design Effect	0.52	0.04	0.44	0.60	0.82	146	20
BAS-GPS	Uncorrected	0.51	0.04	0.42	0.59		146	

the 95% confidence level, the sample is balanced. Notice the absence of gender differences by registration type (Table 5).

Income. Although income measures are more time sensitive than basic demographic variables, the importance of disparities between migrants and nonmigrants warrants attention even though a six-month gap between the two studies makes direct cross-study comparisons questionable.

As with age, mean income estimates—measured as total monthly household income/number of people who depend on this income, in current yuan—reveal very substantial differences across groups. Furthermore, when design effects are accounted for, income disparities widen: not surprisingly, migrant households are the poorest group, with a monthly income of 912 yuan per person (965 when uncorrected for design effect), in contrast to movers (1208 RMB), while correctly registered households stand in between (990 RMB) (Table 6).

Expectation of the impact of the WTO on the Chinese political system. The last example focuses on a time-sensitive attitudinal variable: the likely impact of China's entry into the World Trade Organization (WTO). In December 2001, barely a month after China's formal accession at the Doha conference, BAS respondents were asked to evaluate the current impact of membership on China's structural economic *and* political reforms. The same questions were asked in July 2002, when the initial euphoria of successful negotiations had passed. It is therefore impossible to distinguish how much of the observed cross-sample variation reflects differences in coverage as opposed to actual changes in public opinion.

The results are quite instructive. First, the GPS-BAS respondents are less sanguine about the effect of WTO membership on political reform than in the previous sample. More important, migrants and movers stand once again on opposite ends of the distribution. The affluent are more willing to draw the political implications of membership than other groups, even though the propensity to ascribe political influence

Table 6 Cross-sample comparisons of respondents' income (BAS vs. BAS-GPS)

Study	*Estimation Procedure*	*Mean*	*Standard Error*	*95% Conf. Interval*		*Deff*	*N*	*PSUs*
HOUSEHOLD PERCAPITA INCOME								
-All Respondents Combined								
BAS	Corrected for Design Effect	983	54.17	874	1092	1.39	615	46
BAS	Uncorrected	983	45.96	893	1073		615	
-All Respondents Combined								
BAS-GPS	Corrected for Design Effect	990	83.08	818	1161	3.82	633	25
BAS-GPS	Uncorrected	981	40.33	902	1060		633	
-Migrants								
BAS-GPS	Corrected for design Effect	912	113.62	673	1150	1.69	149	19
BAS-GPS	Uncorrected	963	93.78	778	1148		149	
-Movers								
BAS-GPS	Corrected for Design Effect	1208	121.33	954	1462	1.36	125	20
BAS-GPS	Uncorrected	1148	98.78	952	1343		125	
-Correct Household Registration								
BAS-GPS	Corrected for Design Effect	953	123.14	697	1208	5.15	359	23
BAS-GPS	Uncorrected	931	48.41	836	1026		359	

to the WTO is highly correlated with the respondent's view of the economic impact of membership (Table 7).

Notice that nonresponse patterns differ across studies. In BAS, only 1% of the respondents indicated "don't know" or refused to give an answer to the question on the political impact of membership, in contrast to 10% in the GPS sample (Fig. 7). In BAS-GPS, item nonresponse is very high among migrants: they account for 51% of the total nonresponse. Registered respondents—movers or not—have a high response rate (92%), but the number dips to 72% for migrants. Similar nonresponse patterns apply to the attitudes about the impact of the WTO on structural economic reforms.

4.3 *Multivariate Application*

The combination of differentiated patterns of nonresponse and apparent differences of substantive opinion across groups begs the question of the effect of the GPS design in multivariate analysis.

Since nonresponse varies by household registration status, we first model the respondent's answer as a Heckman sample-selection problem (Table 8). The selection equation simply accounts for the household registration status of the respondent. We use three dummy variables and drop the constant term for identification. The outcome equation models the response as a linear function of demographic variables (age, gender, levels of education), the respondent's job type[13] (including housewives and unemployed), and his registration status. Since preliminary analyses indicate a strong relationship between the perceived impacts of WTO membership on economic and political reforms, the

[13]Since job types are mutually exclusive, each respondent is coded 1 for one of the dummy employment variables. Therefore the model has no constant term.

Table 7 Cross-sample comparisons of respondents views on the political impact of WTO entry (BAS vs. BAS-GPS)

Study	*Estimation Procedure*	*Mean*	*Standard Error*	*95% Conf. Interval*		*Deff*	*N*	*PSUs*
Political impact of WTO [0-10]								
BAS	Corrected for Design Effect	7.04	0.35	6.33	7.76	0.82	613	46
BAS	Uncorrected (completed interviews)	7.04	0.39	6.27	7.81		613	
-All Respondents Combined								
BAS-GPS	Corrected for Design Effect	6.22	0.15	5.9	6.5	1.98	657	25
BAS-GPS	Uncorrected (completed interviews)	6.18	0.11	6.0	6.4		657	
-Migrants								
BAS-GPS	Corrected for Design Effect	5.78	0.33	5.1	6.5	1.86	148	20
BAS-GPS	Uncorrected (completed interviews)	5.70	0.24	5.2	6.2		148	
-Correct Household Registration								
BAS-GPS	Corrected for Design Effect	6.35	0.22	5.9	6.8	2.38	372	23
BAS-GPS	Uncorrected (completed interviews)	6.30	0.14	6.0	6.6		372	
-Movers								
BAS-GPS	Corrected for Design Effect	6.47	0.23	6.0	7.0	1.16	137	20
BAS-GPS	Uncorrected (completed interviews)	6.37	0.21	5.9	6.8		137	

respondent's answer on economic reform is included, along with her or his household's per capita income.

Model 1 confirms that nonresponse is very likely among migrants, but that these differences do not affect the substantive question of interest: ρ is very small and insignificant. The biggest substantive impact remains one's views of the economic impact of WTO membership. Clearly those who believe that membership is affecting China's economic reforms also draw the implication that WTO entry is affecting political reforms as well. The results are invariant to age and education, but not gender, though the impact is modest (the impact of "female" is 0.4 on a 0–10 scale).

Blue-collar and skilled workers—but not farmers—expect political reforms after WTO membership. Cadres, on the other hand, are deeply skeptical that the economic impact of membership is associated with political change, which is likely to reflect the awareness among officials that key policy makers hope to maintain the strategy of moving China along the path of drastic economic reforms without undermining the political status quo of CCP dominance. Notice that registration status is unimportant: once employment type, gender, and views on the economic impact of WTO entry are accounted for, migrants and nonmigrants have similar responses.

Since there is no measurable selection bias by household registration type, Models 2 and 3 test similar specifications without the Heckman treatment. Model 3 replicates the core findings of Model 1 but without sample selection. The substantive impact and levels of significance for most coefficients are virtually identical, as one would expect when ρ is negligible.

Model 2 tests a counterfactual experiment: what would our results look like had the GPS study excluded migrants and movers? This approach amounts to estimating Model 3 solely for the group of 331 respondents who actually reside in the neighborhood where they are formally registered. The results show that even though registration status does not

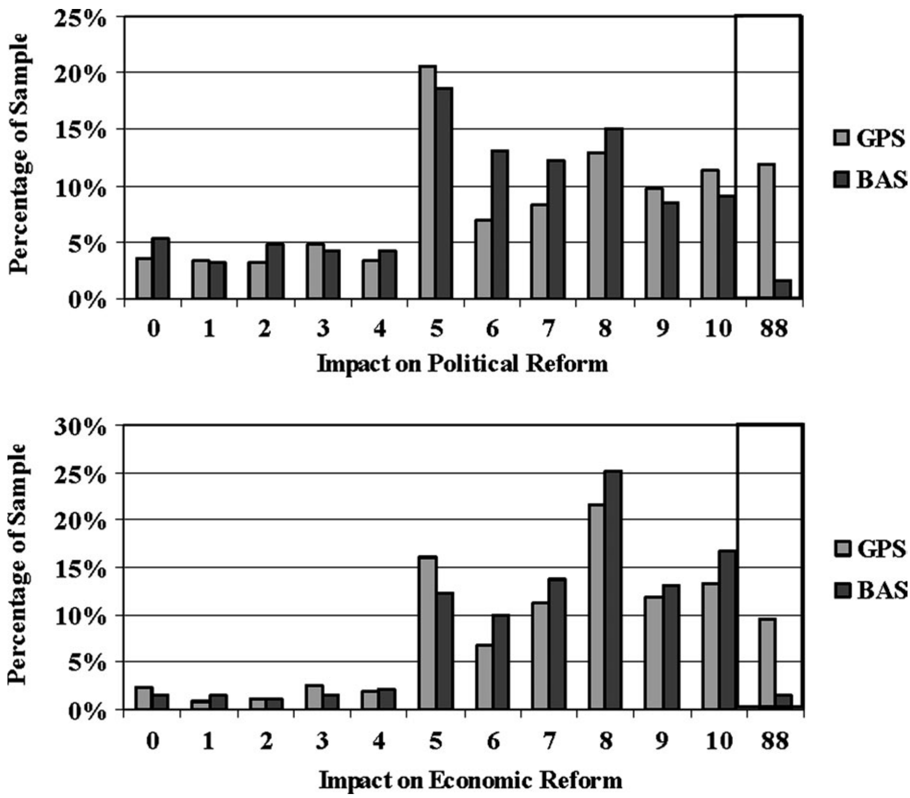


Fig. 7 Respondents assessment of the impact of WTO membership on economic and political reforms, BAS 2001 versus BAS-GPS [0-10 scale] Note: 88 denotes "Don't Know" or "No Answer"

directly impact a respondent's views on the political impact of WTO membership, a sampling strategy that does not ensure proper coverage of migrants or fails to reach one of the most dynamic segments of the population (movers) produces seriously misleading results: Model 2 exaggerates the gender bias ($\beta = .68$ versus .41) and fails to account for the expectation among blue-collar workers, technicians, and teachers that the WTO will lead to political reforms. In Model 2, these coefficients are much smaller than in the full sample and have large standard errors.[14]

5 Conclusions

Social scientists have yet to take advantage of the technological leap that the GPS represents. The theoretical justifications for why spatial samples can be equal probability samples are straightforward, but until recently the inability to draw conveniently small units that could be located, surveyed, and enumerated in a cost-effective manner restricted the use of spatial sampling to the natural and physical sciences. The GPS revolution has solved this technological hurdle: spatial social science surveys have become a practical reality.

[14]However, the lack of statistical significance may be an artifact of estimating Model 2 on the subset of only 331 respondents.

Table 8 Multivariate analysis of the impact of WTO membership on China's political reforms (BAS-GPS, 2002)

		Heckman selection model		*Regression*			
		MODEL 1		*MODEL 2*		*MODEL 3*	
Number of obs.			645		331		564
Censored obs.			81		–		–
Uncensored obs.			564		–		–
				F(15, 22)	998	F(17, 24)	576.7
Prob > F					0.00		0.000
R^2					0.88		0.88
Log likelihood:			−15600000				
Wald $\chi^2_{(17)}$			8776				
Prob > χ^2			0.000				
Number of clusters (PSU)			25		23		25
WTO IMPACT ON POLITICAL REFORM [0-10 scale]							
		Coef.	RSE	Coef.	RSE	Coef.	RSE
Demographic Variables							
Age	[18–66]	−0.01	0.01	−0.01	0.01	−0.01	0.01
Education (0–18)	[0–18]	0.03	0.04	0.04	0.06	0.03	0.04
Gender (1=female)	[dummy]	0.41	0.19**	0.68	0.24***	0.41	0.19**
Job Type	[dummy]						
Entreprenuer	[dummy]	1.17	0.80	1.12	0.88	1.19	0.81
Worker	[dummy]	1.74	0.84**	0.68	0.99	1.75	0.85**
Famrer	[dummy]	0.76	0.71	−0.11	0.81	0.78	0.73
Employee	[dummy]	1.09	0.82	0.19	0.95	1.10	0.83
Teacher	[dummy]	1.71	0.89**	0.41	0.95	1.73	0.90*
Technician	[dummy]	1.60	0.90*	0.58	0.88	1.61	0.91*
Cadre	[dummy]	0.40	1.01	0.29	0.97	0.42	1.03
Housework	[dummy]	−0.42	0.66	−0.38	0.97	−0.42	0.67
Unemployed	[dummy]	0.91	0.70	1.32	0.72*	0.91	0.71
Retired	[dummy]	0.25	0.35	0.56	0.58	0.25	0.35
Economy							
WTO's impact on economy	[0-10]	0.61	0.06***	0.70	0.09***	0.61	0.06***
Per Capita Family Income	[0-10000]	0.00	0.00	0.00	0.00***	0.00	0.00
Registration Status							
Migrant	[dummy]	−0.18	0.34	–	–	−0.15	0.35
Mover	[dummy]	0.16	0.32	–	–	0.15	0.32
SELECTION EQUATION							
Migrant	[dummy]	−0.70	0.23***	–	–	–	–
Mover	[dummy]	0.16	0.19	–	–	–	–
Correct Hukou	[dummy]	1.33	0.13***	–	–	–	–
ρ		0.04	0.04	–	–	–	–
Wald test of indepence across equations: ρ = 0							
$\chi^2_{(1)} = 1.17$							
Prob>χ^2 = 0.28							

All models use probability weights and robust standard errors (clustering by PSU)

The results of our experimentation with GPS sampling are very promising. This article has shown that:

- The method can be implemented practically;
- The method successfully replicates core results for the subset of respondents that are typically reached using alternative sampling techniques;
- Spatial sampling vastly improves coverage of the entire underlying population in an environment in which the institutional capabilities of local governments are limited;
- Both univariate and multivariate analyses show that improved coverage produces substantively different results than traditional approaches.

Yet spatial sampling approach is no panacea. Practical difficulties abound, particularly the problem of correctly forecasting sample size at the design stage. While incorrect forecasts do not invalidate the methodology from a theoretical standpoint, they undermine the cost-effectiveness of the approach. To the extent that any data are available, it seems obvious that better models of the spatial distribution of the population would make the methodology more attractive to cost-conscious researchers.

Both our forecast of sample size and the relative measures of size used to draw PSUs by PPS depended to a significant extent on official data. We believe that our methodology properly corrected for the possible problems associated with the use of such information, because the process by which biases are generated is well understood. In many developing countries or extremely challenging environments, one may not enjoy the luxury of any demographic data. For instance, a study of a refugee camp in a war zone cannot feasibly include ex ante measures of size. In this case, the researcher is better off accepting greater uncertainty about final sample size and drawing units completely at random, in one or mores stages, depending on her tolerance for clustering and the cost of travel from one sampling point to the next. Again, as long as respondents in the final sampling units are all interviewed, an equal probability sample is obtained.

References

Anderson, Margo J., and Stephen E. Fienberg. 2001. *Who Counts? The Politics of Census-aking in Contemporary America*. New York: Russell Sage Foundation.

Brick, J. Michael, Joseph Waksberg, Dale Kulp, and Amy Starer. 1995. "Bias In List-Assisted Telephone Samples." *Public Opinion Quarterly* 59:218–235.

Chamberlain, Robert. "Calculating Distances on the Surface of the Earth." (Available from http://www.census.gov/cgi-bin/geo/gisfaq?Q5.1; accessed Nov. 6, 2002.)

Fan, C. Cindy. 2000. "Migration and Gender in China." *China Review:* 423–454.

Goldstein, Alice, Sidney Goldstein, Shenyang Guo. 1991. "Temporary Migrants in Shanghai Households, 1984." *Demography* 28:275–291.

Goldstein, Sidney. 1987. "Forms of Mobility and Their Policy Implications: Thailand and China Compared." *Social Forces* 65:915–942.

Goodkind, Daniel, and Loraine A. West. 2002. "China's Floating Population: Definitions, Data and Recent Findings." *Urban Studies* 39:2237–2251.

Gui, Shixun, Liu Xian. 1992. "Urban Migration in Shanghai, 1950–88: Trends and Characteristics." *Population and Development Review* 18:533–548.

Johnson, D. Gale. 1994. "Effects of Institutions and Policies on Rural Population Growth with Application to China." *Population and Development Review* 20:503–531.

Kish, Leslie. 1965. *Survey Sampling*. New York: Wiley.

Kish, Leslie. 1987. *Statistical Design for Research*. New York: Wiley.

Kviz, Frederick J. 1984. "Bias in a Directory Sample for a Mail Survey of Rural Households." *Public Opinion Quarterly* 48:801–806.

Liang, Zai, Michael J. White. 1996. "Internal Migration in China, 1950–1988." *Demography* 33:375–384.

Redfern, J. V., et al. 2002. “Biases in Estimating Population Size from an Aerial Census: A Case Study in the Kruger National Park, South Africa.” *South African Journal of Science* 9:455–461.

Shen, Mingming. 2001. “Analysis and Forecast of Urban Resident’s Confidence during the Transition to a Market Economy.” In *Reforms, Development and Social Changes,* ed. Shen Mingming. Beijing: Huaxia.

Skerry, Peter. 2000. *Counting on the Census? Race, Group Identity, and the Evasion of Politics*. Washington, DC: Brookings Institution Press.

Solinger, Dorothy J. 1995. “China’s Urban Transients in the Transition from Socialism and the Collapse of the Communist ’Urban Public Goods Regime.’” *Comparative Politics* 27:127–146.

Solinger, Dorothy J. 1999. *Contesting Citizenship in Urban China: Peasant Migrants, the State and the Logic of the Market*. University of California Press.

SSB (State Statistical Bureau of China). 1997. *Materials on the 1% Population Sample of China*. China Statistics Press.

SSB (State Statistical Bureau of China). 2002. *China Population by Township*. China Statistics Press.

Steenkamp, H. A., C. J. Van Aardt. 2001. *Estimated Undercount of the White Population during the 1996 Population Census for the Randburg Magisterial District*. Pretoria: Bureau of Market Research, University of South Africa.

Subcommittee on Census and Population [U.S. Congress: House Committee on Post Office and Civil Service]. 1987. *Undercount Problem in Los Angeles: Counting Homeless in Decennial Census*. Hearing before the Subcommittee on Census and Population of the Committee on Post Office and Civil Service, House of Representatives, 100th Congress, first session, April.

Tull, Donald S., and Gerald S. Albaum. 1977. “Bias in Random Digit Dialed Surveys.” *Public Opinion Quarterly* 41:389–396.

USCMB (U.S. Census Monitoring Board). 2000. *Estimated Census 2000 Undercount by State*. (Available from http://purl.access.gpo.gov/GPO/LPS30212.)

Wu, Xiaogang, and Donald J. Treiman. 2004. “The Household Registration System and Social Stratification in China: 1955–1996.” *Demography* 41:363–384.

Yang, Xiushi. 1993. “Household Registration, Economic Reform and Migration.” *International Migration Review* 27:796–818.

Zhang, Li. 2001. *Strangers in the City: Reconfigurations of Space, Power, and Social Networks within China’s Floating Population*. Stanford University Press.