Direct Estimation of Hidden Earnings: Evidence from Russian Administrative Data

Serguey Braguinsky *Carnegie Mellon University*
Sergey Mityakov *Clemson University*
Andrey Liscovich *Harvard University*

Abstract

We employ unique administrative data from Moscow to obtain a direct estimate of hidden incomes. Our approach is based on comparing employer-reported earnings to market values of cars owned by the corresponding individuals and their households. We detect few hidden earnings in most foreign-owned firms and larger firms, especially state-owned enterprises in heavily regulated industries. The same empirical strategy indicates that up to 80 percent of earnings of car owners in the private sector are hidden, especially in smaller companies and industries such as trade and services, where cash flows are easier to manipulate. We also find considerable hidden earnings in government services. Our approach sheds new light on the decline in the gross domestic product (GDP) in Russia after the collapse of communism and subsequent recovery; in particular, we argue that a good deal of these changes might represent changes in income reporting rather than actual changes in GDP.

1. Introduction

Measuring the hidden economy is notoriously difficult. Yet it is essential to know both the magnitude and distribution of hidden activity across sectors. Knowing the magnitude allows us to evaluate to what extent recorded differences and growth in gross domestic product (GDP) per capita reflect real differences and improvement in the standard of living. Knowing the distribution of hiding across sectors of the economy allows us to assess actual distortions in resource allocation

We thank Dennis Carlton, William Dougan, Raymond Fisman, Steven Klepper, Tom Mroz, Ben Pugsley, Peter Thompson, and seminar participants at the National Bureau of Economic Research Productivity, Innovation, and Entrepreneurship Working Group; Carnegie Mellon University; Clemson University; and Florida International University for helpful comments. The usual disclaimer applies.

[*Journal of Law and Economics,* vol. 57 (May 2014)]
 0022-2186/2014/5702-0009$10.00

and devise meaningful economic policies aimed at improving efficiency and fighting corruption.

In this paper, we exploit a rare opportunity, in which unique administrative data from Moscow became publicly available, to obtain direct estimates of income hiding across various industries, the government, and public services. Our approach is based on the idea that for some employer types it might be relatively easy not to report large chunks of workers' earnings, but it is very costly for any individual worker to drive an unregistered vehicle.[1] This difference is key to our identification strategy, as we match employer-level administrative data on wages and salaries to car ownership data to measure hidden earnings.

Overall, for a majority of Moscow's car owners, estimated actual earnings exceed reported earnings by a factor of three or more. We detect few hidden earnings for most foreign-owned firms and larger firms, especially state-owned enterprises (SEOs) in heavily regulated industries. We argue that costs of hiding might be particularly high in those cases, as foreign companies might face legal action in their home countries, whereas large companies and companies in more regulated industries are subject to greater scrutiny by tax authorities in Russia. The same methodology results, however, in an estimate of 80–90 percent of earnings being systematically hidden by many smaller private-sector companies, especially in trade and services, where cash flows are easier to manipulate, and in banking and finance. We also find evidence of large hidden earnings among car owners in various public services, especially law enforcement, public education, and health care. We probe the sensitivity of our estimates to different parameters (including varying across individuals) of the relationship between (true) earnings and the value of the stock of cars, and we find them to be very robust. We also check our findings against some more common-sense measures (such as the distribution of the fraction of prestigious cars in the total stock of cars owned by individuals employed by different categories of employers) and find essentially the same results.

Our strategy allows us to identify hidden earnings only for car owners in Moscow and for the time window of data availability (1999–2003). Still, Moscow is by far the most important economic and political center in Russia, and car owners are disproportionately represented among the elite of its private workforce and government officials. Hence, the policy implications of our findings are immense even if limited to the sample that is the direct object of our study. And while there might have been a trend toward greater transparency of income reporting during the 2000s, our sense is that, qualitatively at least, our findings remain relevant in today's Russia as well as perhaps in some other economies with a widespread culture of hiding. We also compare our measure of the transparency of earnings to the measure of managerial diversion and theft in Russian

[1] Moscow police routinely conduct traffic stops to check vehicles' paperwork. Unregistered vehicles may be impounded and can be recovered only after paying a fine and producing a registration document.

companies estimated by Mironov (2013) and find high levels of consistency between our estimates of hidden earnings and his estimates of tunneling and corporate fraud. Thus, the patterns that we document for the hidden earnings of Moscow's car owners could be indicative of more general tax evasion and hiding in the Russian economy as a whole.

We are not the first to attempt to measure the hidden economy. Previous empirical approaches exploit macroeconomic relationships, such as the share of cash in the money supply (Cagan 1958; Tanzi 1983), electricity consumption (Kaufmann and Kaliberda 1996; Schneider and Enste 2000; Alexeev and Pyle 2003), or multiple indicators aimed at econometrically estimating the hidden activity as an unobserved variable (Frey and Weck-Hanneman 1984; Giles 1999). Other studies use survey data on household incomes and consumption expenditures (Pissarides and Weber 1989; Lyssiotou, Pashardes, and Stengos 2004; Ivanova, Keen, and Klemm 2005; Gorodnichenko, Martinez-Vazquez, and Peter 2009). Yet other studies employ direct evidence from tax audits (Klepper and Nagin 1989; Feinstein 1999).

The approach we take in this paper is novel in a few key respects. First, we use administrative microdata. We are thus able to avoid the potential problem that survey respondents with large underreported incomes might misreport their expenditures or even choose not to participate in surveys at all.[2]

Second, our earnings data come from employers, which allows us to obtain direct evidence of hidden earnings disaggregated by employer type. To the best of our knowledge, our approach produces the first estimates in the literature that measure heterogeneity of hidden earnings separately by ownership, industry, firm size, and type of government employment in a large and representative data set.

Our empirical strategy consists of three steps. First, we select from the universe of all data a representative sample of car owners because we use information about car ownership to identify true (as opposed to generally falsely reported) earnings. Second, we use the data on earnings reported by businesses, which we believe to be more or less truthful, and we exploit the variation in incomes and car values within this subsample to estimate the income elasticity of the demand for the stock of cars. Finally, we use that elasticity of demand coupled with the information about car values in the rest of our sample to measure hidden earnings among other car owners.

Our findings significantly broaden the range of plausible estimates of the size of hidden economies in developing and postcommunist countries (see, for example, Fiege and Urban 2008). Our findings also challenge parts of the conventional wisdom about the magnitude and in some cases even the direction in which resources may be misallocated across different sectors of economic activity

[2] In a recent study, Hurst, Li, and Pugsley (2012) found that in U.S. household surveys the self-employed underreported their true earnings by about 25 percent. In countries where survey collectors are even less trusted to keep survey data confidential, this fraction could be significantly higher.

in such countries. Particularly important from this perspective is the reversal of the conventional size-wage effect that we find after taking into account much larger hiding by smaller businesses. Using data on developed countries, studies have consistently found that larger firms pay higher wages to similar workers than do smaller firms (Brown and Medoff 1989; Abowd, Kramarz, and Margolis 1999), and, more recently, this size-wage effect was found to be even larger in developing countries (see, for example, Strobl and Thornton 2002). While there are reasons that greater efficiency may be correlated with larger firm size in advanced market economies, the opposite pattern could be present in developing and postcommunist countries, as smaller firms tend to operate in less rigidly regulated sectors, while larger firms predominantly populate parts of the economy that are state controlled or subject to cronyism. In line with this, we find that both car values and prestigious car ownership rates monotonically and strongly decrease across the firm-size distribution in our data, even though officially reported earnings strongly increase with firm size.

Another important insight from our examination of the heterogeneity of hiding across different types of employers comes from comparing foreign-owned and private Russian firms. Foreign-owned firms in our data pay on average three to four times higher official wages than do domestic firms. This evidence is in line with past findings (Bloom et al. 2010; Holger, Strobl, and Walsh 2007). But looking at car values (or a fraction of prestigious car values, for that matter), we see only marginal evidence of higher actual earnings for foreign-owned firms. Thus, differences in transparency seem to be a factor of first-order importance when assessing the contribution of foreign ownership and foreign management to productivity in countries with a widespread hiding culture.[3]

In Section 2 we discuss the nature of hidden earnings and derive qualitative predictions about the distribution of hiding across different employer types. Section 3 describes the data and our empirical design. Section 4 presents our findings and discusses their external validity. Section 5 examines robustness of our estimates and their sensitivity to estimations of the demand for cars. Section 6 presents some macroeconomic implications of our findings, and Section 7 concludes. Technical details and additional empirical results are presented in an online appendix.[4]

[3] See Braguinsky and Mityakov (forthcoming) for a more detailed analysis of this particular aspect.

[4] The online appendix presents the details of our sample construction and the procedure used to impute car values. It also contains more detailed comparisons between our sample and the general population in Moscow and Russia as well as the U.S. data (to supplement Section 3.3), the details of the procedure used to estimate the income elasticity of the demand for the stock of cars employing several benchmarks (to supplement Section 3.4), as well as additional robustness checks, including the derivations showing the sensitivity of our estimates to possible downward bias in the income elasticity of the demand for cars and to varying elasticity of the demand for cars, to supplement the argument in Section 5.

2. Hiding and Tax Evasion in Theory and Practice

2.1. Some Anecdotal Direct Evidence of Hidden Earnings

In 2011, the biggest private litigation case in the world so far unfolded in High Court in London, where one exiled former Russian oligarch sued another (current) oligarch for cheating him out of $5.5 billion (Harding 2011). The case drew a lot of attention because, for the first time, prominent Russian businessmen publicly and under oath exposed to the whole world the unsavory details of doing business in Russia, with widespread hiding of transactions behind an elaborate network of opaque intermediaries, side payments for protection, and so on. Fittingly, the plaintiff lost the case for the simple reason that he could not produce a single written document to prove that he had indeed owned the assets stripped from him. As it turned out, all arrangements were based on a handshake and carefully hidden from public view, so in the end it was his word against the word of his alleged expropriator.

This bizarre case and its outcome present a snapshot of contractual relations in Russia in general, with labor contracts being no exception. When a firm hires a worker, it is often agreed by handshake that apart from the official salary, the worker will receive unregistered and unreported compensation (black wages), which can often be several times higher than the "white" component. These black wages range from envelopes with cash handed to workers by the management to more elaborate schemes, an example of which is presented below.

As we wrote this paper, we ourselves were surprised by the magnitude of hidden earnings implied by our estimates for large parts of the private sector. We thus sought opportunities to obtain firsthand evidence of hiding that would be independent of any estimation methodology. Such an opportunity presented itself when a friend with insider knowledge of the employment contracts of a group of middle- and top-level managers in two medium-sized private Moscow banks offered to help. This individual (the source) agreed, on the condition of anonymity, to match the 2003 actual earnings of those managers to the reported data at our disposal (the reported data are described in detail in Section 3).

Table 1 presents the comparison of officially reported and actual earnings in those two banks and the implied fraction of hidden earnings.[5] Among top and middle managers of the relatively larger bank A, reported earnings represent just about 11 percent of true earnings, while among top and middle managers of the relatively smaller bank B, reported earnings represent less than 9 percent of true earnings. The weighted average fraction is almost 90 percent.

In absolute terms, the 79 managers in the two banks earned on average more than the equivalent of the $3,000 per month in 2003, but their official average monthly salary was less than $200, totally inadequate to live on in Moscow in that year. The data in Table 1 are, of course, based on evidence from just two

[5] We report only the bank averages in Table 1 even though the source provided us with information about each manager; the fraction hidden was remarkably uniform across all individuals in a given bank.

Table 1
Direct Evidence on True Earnings from Two Private Banks in 2003

	N	True Earnings	Reported Earnings	Fraction Reported
Bank A	45	46,613	3,737	.113
Bank B	34	25,032	471	.089
Weighted average		37,325	2,332	.101

Note. Values are annual averages. Reported earnings are from the administrative database of incomes, validated by the source. True earnings were provided by the source. Both true and reported earnings are in U.S. dollars converted from ruble values using the market exchange rate.

banks. Significantly, however, the true earnings data provided by the source correspond almost exactly to our estimates of true earnings in Moscow's private banking sector that are derived using car ownership data, which we report in Section 4.3.

2.2. Costs and Benefits of Hiding: Empirical Predictions

The economics of tax compliance and tax evasion have been understood reasonably well by economists since at least the early 1990s (see Andreoni, Erard, and Feinstein [1998] for a survey of this literature). In the case of unreported earnings, both employers and employees benefit because employers do not pay the payroll tax (35–39 percent of the amount of reported earnings during the time window of our estimations), while employees do not pay income tax (12–13 percent for most of our sample during our time frame).[6] In other words, reporting earnings would add about 50 percent to the cost of an employee, compared with using a black-wage scheme. If tax compliance is universal or almost universal, standard analysis suggests that the lion's share of this cost would simply be passed through to the employee (assuming that overall labor supply is sufficiently inelastic). The situation becomes very different, however, if the prevailing practice is that of noncompliance. In such a situation, a firm would not be able to pass the taxes through to its employees, so it has a strong incentive to join the ranks of noncompliers.

Of course, hiding also has its costs, both in terms of potential penalties if audited by the tax authorities and in terms of how costly it is to generate off-book revenues that could be used to pay black wages. To fix ideas, consider a simple framework in which by not reporting a fraction x of actual earnings of its employees, the employer can benefit by the amount equal to the tax rate, t, times the unreported amount, xY^*, where Y^* is the total amount of its employees' earnings (and thus depends on firm size). The cost of not reporting, on the other hand, is a function of industry-specific hiding technology and the expected penalty in case the firm is audited and caught hiding. In line with the extant literature on tax evasion, assume that this expected penalty is an increasing and

[6] Individual income tax returns were filed by almost no one in Russia at the time of our analysis (and still are filed only very sparsely). In essence, any relationship a worker in Russia has with income taxation begins and ends with what the employer reports to be his or her earnings. We thank an anonymous referee for directing our attention to this institutional detail.

convex function of the total amount hidden (which will be the case if the tax authority operates with a limited budget and rationally chooses to audit firms when it can expect to uncover particularly large amounts). The problem of a firm operating in sector s and having size Y^* can then be written as

$$\max_x txY^* - a_s C(xY^*),$$

where a_s denotes the technological cost of hiding (which depends on the sector of the economy in which the firm operates), and C is a convex cost function, which depends on the total amount hidden. We treat the firm's overall size Y^* as given for now.[7]

The first-order condition yields the optimal fraction hidden as

$$x^* = \frac{1}{Y^*} C'\left(\frac{t}{a_s}\right).$$

In particular, this implies that hidden income as a fraction of total income will be larger in smaller companies; larger companies facing a convex cost of tax evasion would prefer to report a larger fraction of their activities. Cross-industry differences are also predictable, with industries having higher cost shifters a_s hiding less of their activities.

For example, retail shops, especially smaller ones, and establishments in the service industry and banking possess access to a variety of relatively cheap hiding technologies (such as unrecorded sales or cash-flow opportunities), so we would generally expect these to have low values of the parameter a_s and higher fractions of hidden earnings compared with other industries. The opposite will be true in industries that sell their output at regulated prices. Employers in such industries will have much more difficulty generating sources of hidden earnings for their employees. We thus predict that fractions of hidden earnings will be higher in industries such as trade, banking, and services than in utilities, transportation, and other regulated industries.

We also expect foreign-owned firms to have especially steep expected penalty functions and high technological costs of hiding. Foreign-owned entities, especially large multinational corporations operating in Russia, must be conscious of their reputation, both in their home countries and worldwide. They can also be subject to litigation and punitive sanctions in their home countries for breaking laws in other countries. Moreover, foreign-owned firms commonly lack the necessary connections to escape the scrutiny or strike informal deals with Russian tax authorities, which makes them particularly easy and lucrative targets for audits. We thus expect foreign-owned firms, especially large ones, to pay wages

[7] In the presence of increasing and convex costs of tax evasion with size, the choice of firm size will, of course, be at least partly driven by tax evasion considerations. We will come back to this issue in Section 4. For now we assume that production technology and not tax evasion considerations dominate the choice of the optimal firm size.

and salaries to their employees in Russia transparently and to have little if any hidden earnings in their employment contracts.

With competitive labor markets and salaries in government service benchmarked against officially reported (as opposed to actual) earnings in the private sector, we also expect to see hidden earnings in public services and government. These, however, tend to be individual rather than employer specific, as they most likely represent illicit fees and bribes. We conjecture that certain areas of government services that are especially prone to corruption, that is, where demand for services is inelastic as there are no close substitutes (for example, law enforcement, public education, and health care), will have more hidden earnings than state-owned manufacturing, communications, services, and so forth.

3. Data and Empirical Design

In the early to mid-1990s, amid official statistics showing sharply declining GDP and plunging real incomes of the vast majority of the population, one could not help but notice the ever-increasing number of expensive foreign cars in the streets of Moscow, St. Petersburg, and other Russian cities. For Russia as a whole, the car ownership rate more than doubled from 58.6 vehicles per 1,000 people in 1990 to 122 vehicles per 1,000 people in 1998 (the last year of the official GDP decline). For Moscow, the corresponding number almost tripled, from 70.6 to 200.4 (Federal State Statistics Service 2003, table 7.3).[8] Part of this explosion in car ownership, especially in the early years, followed from the scarcity of cars in the Soviet Union. Still, new-car purchases could hardly have taken place on such a scale if real earnings were indeed declining sharply, especially given the fact that consumer credit was all but absent until 2004, and the lion's share of vehicle purchases were made using cash.[9] The obvious discrepancy between the trends in officially measured earnings and in cash-financed car ownership motivates our approach.

3.1. Data Sources

The data used in this paper come from three sources. One is the residency registry for the year 2002. It contains the names and addresses of all registered residents of Moscow and can be used to identify members of the same household.

The second source is the 2005 Moscow auto registration database (containing retrospective data on vehicle ownership), which can be matched to the residency database. We use the vehicle's identification number (VIN) to trace its history

[8] Unless explicitly stated otherwise, all macroeconomic data are from the official Federal State Statistics Service, also known as Rosstat.

[9] We could not find official statistical data in Russian sources, but Western estimates put the fraction of cars bought on credit in Russia at about 7.3 percent as late as 2002 (Ernst & Young 2008). Consumer credit started rapidly expanding in 2004 (Bush 2004), and, reflecting this, the fraction of cars bought using credit increased to almost 50 percent by 2007. This has no effect on our analysis, which ends in 2003.

of owners. We use the information about the make, model, and year to impute the market value of the car in a given year.

The third and most unique source consists of administrative databases of incomes containing the universe of reported incomes filed by officially registered employers and other income-generating sources in Moscow for 5 years, from 1999 to 2003. The databases became available around 2004 as a result of a historic episode that placed them in the public domain, whether accidentally or through a leak is not quite clear. In any event, today, more than 10 years after these data were released, first burned on compact discs and then posted on the Internet, they clearly possess historical and research value and have already been employed in this capacity by academic researchers, including the current authors (see for example, Guriev and Rachinsky 2008; Braguinsky 2009; Mironov 2013; Braguinsky and Mityakov, forthcoming).[10] Mironov (2013) and Mironov and Zhuravskaya (2012, pp. 5–6) discuss the authenticity of these data in detail, as did Russian journalists at the time of their release (see, for example, *Forbes Magazine* [Russia] 2004; *Vedomosti* 2004, 2005). We also conducted our own checks of the reported income data sources, in particular comparing official Moscow labor statistics as published by the Russian Federal State Statistics Service (Rosstat) to sample averages obtained from our databases for the same years and found a close match both for all 5 years on average and in any given year, as well as for most sector-year averages.

Even though the data were procured from the public domain, in the sample used for the purposes of this paper, we removed all individual- and employer-identifying information after using it to match the earnings data to car ownership and to retrieve information from open sources about the employer's sector of economic activity and type of ownership. Our data (with individual- and employer-identifying information removed) and our codes are available to anyone who might want to replicate our findings, while the databases themselves can be purchased online.

3.2. *Sample Construction*

In theory, the information contained in the administrative databases of incomes would allow us to match them across years and to the residency and vehicle registration databases. In practice, however, such matching presented a lot of problems because of vastly different formats of records across databases, missing fields, typos, duplicate entries, and the like. To improve the quality of the matching process, we decided to manually verify the initial matching procedure done by automated script. This made it too costly to work with the universe of all the data and forced us to choose a smaller sample.

[10] The data are still available in the public domain free of charge or for a nominal price (for example, as we wrote this paper, the Russian Web site Mos-Inform.com [http://www.mos-inform.com], which has been online for many years, offered these data on three DVDs for just 1,000 rubles, which is less than $30). We do not use the 2004 data set, which is also available, because the spread of consumer credit in Moscow that started that year could interfere with our identification strategy.

We started with 30,000 households (97,141 individuals) randomly selected from the residency database, which had been first purged of poor-quality data to improve the matching process,[11] and then matched individuals in this sample to the income and auto databases using full names, dates of birth, and addresses. We then eliminated commercial vehicles (such as trucks and buses), motorcycles, and households consisting solely of retirees. This left 6,101 households (21,617 individuals) for which we could match at least one member to the auto registration database and to officially reported earnings in each of the five income databases. The latter condition was imposed to ensure that we have 5 consecutive years of officially reported earnings, thus excluding temporary residents and those whose sole source of income was the informal sector in any year between 1999 and 2003.[12] We then manually eliminated duplicate income and auto records while adding nonduplicate earnings and vehicle entries to obtain the total for an individual and his or her household. The ownership and sector of employment were assigned on the basis of the source of the highest reported income.

The classification of employers (income-generating sources) was conducted using unique employer identification numbers and employers' names contained in the administrative databases of incomes. There are a total of 13,263 distinct employers in our sample. Most of the classification had to be performed manually by matching employers' identification numbers, names, and addresses to information from open sources.[13]

Using the earnings data for each employer in the database, we calculated for each individual in our sample his or her relative position in the employer's hierarchy as the percentile in the employer's overall earnings distribution in a given year. Even though large fractions of true earnings may be hidden, the relative position of the employee in the employer's earnings distribution generally reflects where the person is positioned in the employer's hierarchy.[14] We also calculated the size of the employer by counting the number of entries pertaining to its identification number in any given year. For expositional convenience, ruble values were converted to U.S. dollars using average market exchange rates for each year.

[11] In particular, we excluded all names with non-Cyrillic characters at this stage, so that we have only Russian nationals in our sample, including employees of foreign-owned firms.

[12] About 30 percent of passenger car owners in the vehicle registration database did not meet this criterion, which means that they did not continuously reside in Moscow, were not formally registered by their employers, or were self-employed in at least 1 year. Our findings refer to the degree of hiding in formally registered employment only.

[13] There are multiple Web sites offering information about Russian employers, such as their sector of activity and ownership, free or for a nominal price (for example, the Moscow Center for Economic Security [http://www.businessinfo.ru] and Globalstat [http://globalstat.ru]). The information on these Web sites is not always accurate, however, as we found repeatedly when comparing them with our manual classifications. In particular, firms classified as foreign owned are often offshore companies set up using Russian capital, which is a serious issue for our identification strategy.

[14] Even though absolute amounts of earnings could be severely underreported, the ordering of reported earnings preserves the ordering of actual earnings within an employer. Note in particular that if the ordering of reported earnings was at odds with the actual within-firm hierarchy, this would be a clear red flag to the authorities and could prompt an audit.

The total number of person-year observations with nonzero reported earnings in our sample is 56,893, of which 33,307 (58.5 percent) originated in the Russian private sector, 1,320 (2.3 percent) in the foreign-owned sector, and 22,266 (39.1 percent) in the government (federal and Moscow) and state-owned sector. The total number of person-year observations with nonzero car values is 40,281, but about 30 percent of those do not have VINs, mostly because they are older Soviet-made cars that have no VINs. Since such cars had heavily depreciated by the time of our estimation time window (although 30 percent of the total number of cars, their fraction of the total value of cars in our sample is much smaller), we elected to exclude them from our analysis because we could not determine the period during which an individual owned a given car.[15] We also eliminated a very small number of individuals who owned more than five cars in any given year out of concern about the quality of the data in those cases.

To make sure our estimates capture the extent of income underreporting by type of formal employment, we took extra steps to eliminate potential problems from possible income sources unrelated to such employment. First, in addition to requiring at least one member of a car-owning household in our sample to have had an administratively registered employer in all 5 years during 1999–2003, we exclude all those who earned less than the official minimum wage in any year (5–7 percent of observations, depending on the year). We also exclude car owners whose reported earnings exceeded the equivalent of U.S.$100,000 in any year (less than .3 percent of observations). Second, we looked at all sources of income and identified nonwage earnings, such as lottery prizes, income from financial assets, social security and insurance payments, and the like. We exclude from our estimates 1,741 individuals (and their households) whose main source of income came from such nonwage sources. Thus, our analysis pertains mostly to middle-class residents of Moscow whose main source of income is from an employer and who have enough resources to own a car. We discuss the representativeness of this sample in Section 3.3.1 and the external validity of our results in Section 4.4.

3.3. Summary Statistics

3.3.1. Car Owners versus the General Moscow Population

Table 2 presents basic summary statistics for our sample for all years and for 2003 and compares them with Rosstat's official statistical data on Moscow's workforce in 2003. The average annual reported earnings in our 2003 sample

[15] We reran our estimations assuming that cars with missing vehicle identification numbers were owned by individuals in our sample throughout 1999–2003. The basic results were not affected. Another caveat of car ownership in Russia at that time was that vehicles were sometimes sold without formal transfer of the title. The new owner of the vehicle simply received a document from the original owner that authorized him or her to drive the vehicle. We have no means of knowing to what extent this practice may affect our sample, but the practice was rather risky and was largely limited to older and less expensive cars. From the point of view of the focus of our analysis, such de facto transfer of ownership without formal transfer of the title simply introduces some noise uncorrelated with ownership or the sector of employment.

Table 2
Summary Statistics: Earnings

	Administrative Databases of Incomes		Rosstat Data, 2003
	1999–2003	2003	
Mean annual reported earnings (U.S.$)	4,500	6,177	4,158
	(8,280)	(9,671)	
Mean car values (U.S.$)	7,577	9,069	
	(11,452)	(13,865)	
Mean car values/earnings	1.68	1.47	
Mean age (years)	43.1	42.4	40.7
Fraction male	.75	.74	.51
Fractions employed in:			
Private firms	.656	.663	.636
Government and SOEs	.306	.303	.283
Foreign-owned firms	.038	.034	.082
Banking, finance, and insurance	.080	.077	.031
Federal and city government	.058	.058	.035
Education and science	.096	.099	.114
Health care and sports	.055	.057	.077
Construction	.115	.114	.141
Utilities	.025	.025	.041
Transportation and warehousing	.045	.044	.054
Trade and services	.278	.280	.291
Manufacturing	.133	.132	.116
Communications and IT	.040	.040	.029
Other	.075	.074	.042
N	19,882	4,935	5,631,300

Source. Rosstat data on Moscow are from Federal State Statistics Service (2005).
Note. Statistical data on the fraction male and fractions employed in private firms, government and state-owned enterprises (SOEs), and foreign-owned firms are not available for Moscow separately, so averages for the whole Russian Federation are presented. Standard deviations are in parentheses. The sample consists of car owners with nonzero reported earnings in all 5 years, excluding those with five or more cars, those under 18 in 1999 and older than 60 in 2003, those earning less than the legal minimum wage or more than U.S.$100,000 in a given year, and those whose primary reported earnings in any year 1999–2003 came from nonwage sources. IT = information technology.

are 50 percent higher than in the Rosstat data. Car owners are slightly older than the average Moscow worker and are disproportionately male. In Table 3, we see that the average percentile of car owners in the overall earnings distribution in our administrative databases of incomes is .74, with averages for private and government employment being .75 and .70, respectively. Thus, an average car owner in both private-sector and government employment is positioned much higher in the hierarchy of his or her employer than an average worker. The median earner is located at the 80th percentile of the earnings distribution, while those above the 90th percentile belong to the top 1 percent of all income earners in Moscow.

Our main variables of interest are ownership and sector of economic activity of the employer, and Table 2 shows that the sample is quite representative in this regard. The fraction of employees of foreign-owned firms is lower in our

Table 3
Summary Statistics: Car Owners in the Overall Earnings Distribution

	All Car Owners	Private Employment	Government Employment
Mean	.74	.75	.70
	(.24)	(.24)	(.24)
Percentile:			
5th	.25	.27	.23
10th	.37	.38	.32
25th	.58	.60	.54
50th	.80	.81	.76
75th	.94	.94	.91
90th	.99	1.00	.97
95th	1.00	1.00	.99

Note. The sample consists of car owners with nonzero reported earnings in all 5 years, excluding those with five or more cars, those under 18 in 1999 and older than 60 in 2003, those earning less than the legal minimum wage or more than U.S.$100,000 in a given year, and those whose primary reported earnings in any year 1999–2003 came from nonwage sources. Percentiles are the percentile of car owners in the overall earnings distribution in all databases at the mean and at the corresponding percentiles. For example, a car owner at the fifth percentile in the sample is at the 25th percentile in the earnings distribution overall. Standard deviations are in parentheses.

sample than in the Rosstat data, but this is because the Rosstat data include offshore Russian firms in this category.

The fractions of individuals employed in different sectors of economic activity in our sample and in the Rosstat data are for the most part also similar. Deviations from the Rosstat data are easy to explain—for example, it is not surprising that there are relatively more car owners among those employed in banking and finance, federal and city government, and communications and information technology (IT) because those sectors would normally be associated with higher than average (actual) earnings. Similarly, one would normally associate sectors such as construction, utilities, transportation, and warehousing with lower (actual) earnings, and these occupations are indeed somewhat less well represented in the sample of car owners than in the Rosstat data. Overall, Table 2 shows that all types of ownership and sectors of the economy are broadly represented in our sample of car owners.

3.3.2. Comparison with the U.S. Data

Table 4 presents basic statistics on reported earnings and car values in our sample. Average car values exceed average annual earnings for all years and in each year. The household-level data show very similar patterns. On average, car values exceed annual incomes by 68 percent at the individual level (43 percent at the household level), and while there is some tendency for the ratios of car values to reported income to decrease over the years, the total value of cars owned by individuals or households always exceeds annual earnings.

To put those numbers in perspective, we compare data from our sample to U.S. data, where hidden earnings should be negligible (at least compared with

Table 4
Average Earnings and Car Values

	N	Earnings Mean	Earnings SD	Car Value Mean	Car Value SD	Car Value/ Earnings
Individual level:						
All	19,890	4,499	8,279	7,575	11,450	1.68
1999	2,991	2,542	5,913	6,526	10,938	2.57
2000	3,470	3,043	6,005	6,569	10,162	2.16
2001	4,076	4,483	8,665	7,519	11,370	1.68
2002	4,483	5,070	8,621	7,800	11,012	1.54
2003	4,870	6,224	9,716	8,778	12,875	1.41
Household level:						
All	20,671	6,410	9,472	9,167	13,466	1.43
1999	3,313	3,846	6,847	7,747	12,704	2.01
2000	3,759	4,616	7,375	8,000	12,146	1.73
2001	4,252	6,272	9,499	9,238	13,685	1.47
2002	4,589	7,214	9,796	9,677	13,307	1.34
2003	4,758	8,962	11,257	10,521	14,708	1.17

Note. The sample consists of car owners with nonzero reported earnings in all 5 years, excluding those with five or more cars, those under 18 in 1999 and older than 60 in 2003, those earning less than the legal minimum wage or more than U.S.$100,000 in a given year, and those whose primary reported earnings in any year 1999–2003 came from nonwage sources..

our data), obtained using the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). Since the NLSY data do not distinguish between cars owned by respondents and their spouses, we use the data on total net family income in the previous calendar year and market value of vehicles to compare with household-level data for Moscow. In order to make samples more comparable, we take the following steps. For the sample of Moscow car owners we limit the years to 2000 and 2003, for which the administrative databases of incomes contain codes allowing us to distinguish earnings from primary employment (wages and salaries) from other income sources. We also require Moscow car owners to be in the same age bracket as the NLSY79 respondents (38–46 years old in 2003) and impose the condition that the total household income be equal to the wage earnings of the car owner (so that the car owner is the sole wage earner in the family). Similarly, in the NLSY data we require total net family income to be equal to total income from the wages and salary of the respondent and limit the sample to respondents who resided in central cities in metropolitan statistical areas (MSAs) to make them more comparable to Muscovites in terms of infrastructure conditions.[16]

Table 5 presents these comparisons. We find that in Moscow the ratios of car values to earnings range from 1.59 for single-member households to 1.86 for

[16] Moscow is deservedly famous for its highly efficient rapid transit system. If anything, this could lead Muscovites to demand fewer cars, which would bias the results against our findings. Unfortunately, we cannot use proximity to rapid transit systems as an instrument for car ownership because areas close to rapid transit also have more expensive and prestigious housing and thus tend to be populated by more car owners and/or car owners with more expensive cars.

Table 5
Earnings of Households with Cars: Moscow Data and NLSY79 Data

		Wage Earnings		Car Value		
	N	Mean	SD	Mean	SD	Means Ratio
Single-member households:						
Moscow	231	5,787	11,978	9,197	14,344	1.59
NLSY79	351	32,568	18,436	9,361	12,031	.29
Households of two or more:						
Moscow	976	5,221	9,691	9,714	15,011	1.86
NLSY79	809	32,284	19,071	9,830	11,382	.30

Note. Moscow data are from households with wage earners as sole earners in 2000 and 2003 pooled together, truncated at U.S.$100,000, and 38–46 years old in 2003. The National Longitudinal Survey of Youth (NLSY) data are from 1979 cohort respondents to the 2000 and 2004 surveys pooled together, residing in central cities in metropolitan statistical areas with total (truncated) net family income in the past calendar year equal to total (truncated) respondents' income from wages and salary in the past calendar year.

households with dependents. In the NLSY data, the corresponding ratios are below .3. Most strikingly, earnings in the NLSY data in Table 5 are about five to six times higher than reported earnings in Moscow, while car values are quite similar. In other words, if Moscow earnings data are reported accurately, it must be true that Russian wage earners could afford cars of similar value to those of their counterparts in the United States while having earning flows that are five to six times lower (and no access to consumer credit).

3.3.3. Comparison across Characteristics of Employers

To verify our intuition about relative costs of tax evasion depending on firm characteristics, in Table 6 we present the 2003 data on earnings and market values of cars for full-time wage earners disaggregated by several employer characteristics. Table 6 reveals several suggestive patterns consistent with our empirical hypotheses discussed in Section 2.2. First, we argued that foreign-owned companies might be more transparent since they are likely to have higher costs of tax evasion than domestic firms, for example, because of the possibility of legal action in their home countries or lack of connections in Russia. Indeed comparing wage earners in private Russian firms and foreign-owned firms, we see that employees of foreign-owned firms officially receive wages and salaries that are more than four times higher than those of employees of private Russian firms, but their car values are just marginally (5 percent) higher, a difference that is is not statistically significant at conventional levels using a double-sided *t*-test. Workers employed by the government and SOEs have reported earnings that are 26 percent higher than those of workers employed in the private sector, but their car values are 30 percent lower.[17]

[17] One may hypothesize that foreign-owned firms and/or government agencies may provide company cars to their employees, but this practice is not widespread in Russia and in any event cannot extend beyond the very top of the employer hierarchy. To gauge how important this factor may be,

Table 6
Car Values and Earnings in 2003 by Employer Type

	Mean			Mean/SD	
	Earnings	Car Value	Earnings/ Car Value	Earnings	Car Value
Private Russian firms:	5,180	9,566	1.85	.59	.68
Wholesale and retail trade	3,200	9,599	3.00	.62	.75
Mass media, IT, and utilities	9,822	8,346	.85	.80	.86
Firm size below median	3,352	10,777	3.21	.50	.71
Firm size above median	7,024	8,344	1.19	.69	.66
Government and SOEs:	6,513	6,612	1.02	.82	.73
Law enforcement, secondary education, and health care	3,683	6,207	1.69	1.18	.70
Other	7,440	6,744	.91	.84	.74
Foreign-owned firms	21,543	10,017	.46	1.11	.98

Note. The sample is limited to 2003 wage earners, ages 18–61, with earnings above the minimum wage and below U.S.$100,000 and no more than five cars per individual. IT = information technology; SOE = state-owned enterprise.

Table 6 also suggests that income hiding might be more prevalent in industries where cash flows are easier to manipulate (such as wholesale and retail trade) than in more visible or more capital-intensive and regulated industries (such as media, utilities, and IT) even within the private sector. In terms of firm size, we find that reported employee compensation is much higher in larger companies, but car values are much higher in smaller companies, which suggests the opposite income ranking.[18]

Finally, there are considerable differences within government employment. Car owners employed in law enforcement, public health care, and secondary education on average earn 50 percent or less than employees in other branches of government and in SOEs. Their car values are, however, statistically indistinguishable. Workers in law enforcement, health care, and education also have the highest mean-to-standard-deviation ratios in earnings of the employer types in Table 6 (which suggests that they are officially paid according to more egalitarian principles than other categories of employees) but one of the lowest mean-to-standard-deviation ratios in car values, which suggests greater inequality in real incomes compared with other categories (likely because of idiosyncratic side payments or corruption opportunities). It is also noteworthy that employees of foreign-owned firms are paid much more equally than employees of private Russian firms and even most government employees, judging both by the mean-to-standard-deviation ratio of reported earnings and (especially) by the corre-

we compiled the analogue of Table 6 while excluding the top 5 and then also the top 10 percent of workers in the earnings distribution of their employers. The results (not shown) were very similar to those in Table 6.

[18] Note that we exclude from our sample commercial grade vehicles, vans, and trucks, which could be used as capital rather than consumption goods by small businesses.

sponding ratio of car values. Next we evaluate those patterns more rigorously through the lens of an empirical model.

3.4. Empirical Model

We start from the assumption in Section 2 that employers report only a fraction of the earnings of their employees and that this fraction might be systematically related to differences in firm size and sector-specific technologies for hiding. In our data, however, we can observe these systematic differences across employers only by comparing individual realizations of employees' car values with their reported earnings. Here we develop an estimation framework that allows us to test our predictions about systematic differences in the transparency of employment contracts by employer ownership, size, and industry while using employee-level data on reported earnings and car values.

Let employee i's earnings at time t be reported in the amount $E_{it}^{\mathrm{R}} = \Gamma_{it} E_{it}^{*}$, where E_{it}^{*} is true earnings and Γ_{it} is the fraction reported. This fraction may depend on a range of individual- or household-specific characteristics and on time because of, for example, institutional changes. Most crucial, however, Γ_{it} depends on the costs and benefits of hiding that vary with employer characteristics. Hence, we assume that

$$\ln E_{it}^{\mathrm{R}} = \ln E_{it}^{*} + \beta' \boldsymbol{S}_{it} + \gamma' \boldsymbol{X}_{it}^{(1)} + \varphi_1(t) + \eta_{it}, \tag{1}$$

where $\boldsymbol{X}_{it}^{(1)}$ is the vector of individual characteristics and $\boldsymbol{S}_{it}$ is the vector of employers' characteristics (such as ownership type, size, and industry). The primary focus of our empirical analysis is the vector of coefficient β, which measures average income hiding associated with different characteristics of employers $\boldsymbol{S}$. The more negative coefficient β_k is, the larger is the fraction of hidden earnings of total earnings among individuals employed in the category of employers possessing characteristic k.

Our main identifying assumption is that while the fraction of reported income depends on the sector of employment, the demand for the stock of cars has the same functional form in all sectors. In particular, we consider the following car stock demand equation:

$$\ln C_{it} = \lambda \ln E_{it}^{*} + \gamma_2' \boldsymbol{X}_{it}^{(2)} + \varphi_2(t) + u_{it}. \tag{2}$$

That is, the demand for cars depends on actual earnings E_{it}^{*}, individual characteristics $\boldsymbol{X}_{it}^{(2)}$, time effects $\varphi_2(t)$, and an individual- and time-specific disturbance term u_{it}, $E[u_{it}] = 0$.

It is worth emphasizing that equation (2) does not mean that car demand is independent of the sector of employment $\boldsymbol{S}$. The sector of employment may (and very likely will) affect the demand for cars, but we assume that this happens only indirectly, through the channel of true earnings and individual characteristics that may be different in different sectors. This assumption is discussed below, but first we lay out the rest of our estimation strategy.

Substituting for $\ln E^*$ from equation (1) into equation (2), we get

$$\ln C_{it} = \lambda \ln E_{it}^{\mathrm{R}} - \lambda\beta'\boldsymbol{S}_{it} - \gamma'\boldsymbol{X}_{it} - \varphi(t) - \varepsilon_{it}. \tag{3}$$

Here $\gamma = \lambda\gamma_1 - \gamma_2$ and $\varphi(t) = \lambda\varphi_1(t) - \varphi_2(t)$ measure the combined effect of observables and time effects,[19] $\varepsilon_{it} = \lambda\eta_{it} - u_{it}$ is the combined disturbance term, and $\boldsymbol{X}_{it}$ combines individual characteristics from $\boldsymbol{X}_{it}^{(1)}$ and $\boldsymbol{X}_{it}^{(2)}$.

Unfortunately, we cannot estimate equation (3) consistently, because reported income E_{it}^{R} is correlated with the error term ε (which contains disturbance η from the income underreporting equation [1]). For the same reason, the regression with car values C_{it} as the explanatory variable would also produce biased estimates. However, if we were able to consistently estimate the income elasticity of demand for the stock of cars λ, then it would be possible to estimate sector-specific income hiding from the following regression:

$$\ln E_{it}^{R} - (1/\lambda)\ln C_{it} = \beta'\boldsymbol{S}_{it} + \gamma'\boldsymbol{X}_{it} + \varphi(t) + \varepsilon_{it}. \tag{4}$$

We cannot directly use car demand equation (2) to estimate λ since actual incomes E^* are not observed. However, parameter λ can be estimated from other data sets in which falsifying reported earnings is not that much of a problem. One such opportunity is presented by the NLSY79 data. While it can raise questions about applicability of the estimation results to Russia, evidence presented in Table 4 and especially in Table 5 suggests that preferences for cars may not be so different between Muscovites and residents of urban areas in the United States. The second approach, which does not rely on the assumption of cross-country similarity in the demand for cars, is based on the subsample of Moscow car owners (all Russian nationals) employed by a specific category of employer that is likely to report earnings of its employees transparently, namely, foreign-owned firms.

We first estimate the elasticity of the demand for the stock of cars using the subsample of car owners employed by foreign-owned firms. To increase the number of available observations, we add to the original data all car owners employed by foreign-owned firms identified as such in our full sample. Since we do not have household data for those additional observations, we estimate the income elasticity of the demand for cars using only individual earnings and car values. Preliminary analysis suggests that some foreign-owned firms, specifically smaller ones, may not be quite transparent in reporting their employees' earnings either, so we use the subsample of car owners employed by for-profit foreign-owned firms with at least 100 employees in all years (see Section A5 of the online appendix for more details). The estimation equation is equation (2), with earnings assumed to be truthfully reported in this category of employers.[20]

[19] This means that we cannot identify separately the effects of individual characteristics and time effects on income hiding as opposed to the demand for cars, unless we impose some exclusion restrictions. We identify sectoral income hiding by assuming that sectoral dummies are not part of the car stock demand equation.

[20] If there is some residual hiding even in this subsample, our estimates of hiding among the rest of Moscow car owners should be interpreted as a lower bound.

The parameter of interest is λ, which measures the income elasticity of the demand for the stock of cars. Estimating equation (2) on this subsample of car owners produces a value of λ in the range from .3 to .4, regardless of whether industry dummies are included among the controls, and is in line with our main identifying assumption that the demand for cars depends on the sector of employment only through income.

We next conducted the same estimations using the NLSY79 household-level data, matching the industries in those data to our industry classification in the Moscow data, and found a value of λ equal to .74 (see Section A5 and Table A5.4 of the online appendix for details of the estimation procedure and the results). Once again, the estimated value of λ is the same regardless of whether industry dummies are included as controls.[21] Finally, we constructed an analogue of individual-level data using the NLSY79 single-member households and estimated equation (2) to obtain a value of λ equal to .5.

In Section 4 we use the value of $\lambda = .74$ obtained from household-level NLSY79 data as our benchmark parameter value when using the data aggregated to household level and the value of $\lambda = .35$ obtained from individual data on car owners employed by relatively large foreign-owned firms in Moscow when using individual-level data. This allows us to compare the estimates of hidden earnings using values of λ from very different data sources. In order to check how sensitive our estimates are to choosing a different value of λ, we also use the value of $\lambda = .5$ obtained from single-member NLSY79 households. It turns out that our main findings concerning the magnitude of hiding are quite robust to a rather broad range of the values of λ as well as to specifications in which income elasticity could vary with income. This is discussed in more detail in Section 5.

4. Results

Here we present the results of estimating employment-specific hidden earnings using equation (4), utilizing household-level and individual-level data and the corresponding values of the parameter λ, as explained above. In individual-level regressions we control for observable characteristics such as age, gender, ethnicity, and the number of members in the household to account for differences in family composition and preferences. We also include year dummies and the percentile of earnings of a given individual in the earnings distribution of his or her employer and its squared term.

For estimations using household-level data, we constructed household-level analogues of the variables capturing ownership and sector of employment as follows. We first assigned each individual-year observation a weight equal to the

[21] In the National Longitudinal Survey of Youth data, none of the industry dummies were statistically or economically significant. An *F*-test failed to reject the hypothesis that the coefficients are jointly equal to zero.

share of that individual in the total (reported) household earnings in that year. For each household-year observation, we then calculated averages of sectoral and ownership dummies using the above weights. The household-level variables capturing employer size and the percentile in the employer's earnings distribution were constructed similarly. We used the simple mean across all members of the household to construct household average age, gender, and ethnicity. We also perform household-level estimations below using the primary earner's characteristics instead of household-average characteristics as controls, and the results are very similar.

4.1. Ownership and Employer Size Effects

We first use equation (4) to test our predictions about the impact of firm ownership and size on the transparency of employment contracts. The regression includes year and industry dummies. To focus on hidden earnings (as opposed to income from bribes and corruption in general), we exclude car owners employed in federal and city government, law enforcement, state-run education, science, and health care. We examine those sectors in more detail below.

Table 7 presents estimation results for the household-level data (with $\lambda = .74$) and for the data using individual incomes and car values with the income elasticity parameters $\lambda = .35$ and $\lambda = .5$. The coefficients on ownership and firm size strongly support the predictions in Section 2. In particular, foreign-owned firms have much higher reported earnings for employees than state-owned firms or the omitted category (private domestic firms). The effect is robust across all specifications. The magnitude of the difference is quite large—the estimates for household-level data imply that foreign-owned firms report earnings that are on average 3.74 times ($= \exp(1.318)$) higher for employees with the same car values than those of private Russian firms. This gap is even larger in individual-level regressions. In contrast, there is no significant difference between the overall transparency of employment contracts of private and state-owned enterprises.

Similarly, we find much more transparent employment contracts in larger firms than in smaller firms, which confirms our predictions. Once again, the magnitude of the differential between small and large firms is important. Point estimates suggest that doubling employer size increases reported incomes for the same car values by 25 percent in the household-level regressions and by 30–35 percent in the individual-level regressions. Firms at the 95th percentile of the firm size distribution (FSD) are estimated to report 1.9 times more earnings for the same car values than do firms at the fifth percentile of FSD according to household-level estimates and 2.2–2.6 times more according to the individual-level data.

4.2. Relative Hidden Earnings by Sector of Economic Activity

We next test the predictions about the scope of hidden earnings in different sectors resulting from possible differences in technology of hiding across in-

Table 7
Estimates of Regression (4): Ownership and Size Effects

	Household Level, $\lambda = .74$	Individual Level $\lambda = .35$	Individual Level $\lambda = .50$
Ownership:			
State	.075	.211	.201^{+}
	(.087)	(.154)	(.111)
Foreign	1.318**	1.436**	1.550**
	(.122)	(.203)	(.148)
Log number of employees	.245**	.357**	.304**
	(.014)	(.024)	(.017)
Individual (household) characteristics:			
Percentile in EED	6.000**	7.450**	5.883**
	(.536)	(.780)	(.564)
Percentile in EED squared	−3.708**	−5.772**	−4.076**
	(.420)	(.629)	(.455)
Household members	.008	.038	.019
	(.019)	(.031)	(.023)
Age in 2003	.005	.052**	.039**
	(.003)	(.005)	(.004)
Male	−.297*	−.092	−.088
	(.131)	(.122)	(.086)
Year dummies:			
2000	.142**	.118**	.144**
	(.030)	(.049)	(.037)
2001	.341**	.227**	.342**
	(.037)	(.060)	(.045)
2002	.431**	.309**	.474**
	(.040)	(.068)	(.050)
2003	.503**	.260**	.495**
	(.043)	(.073)	(.053)
Constant	−5.048**	−22.707**	−14.851**
	(1.144)	(.718)	(.512)
N	12,697	15,754	15,754
Adjusted R^2	.167	.108	.146

Note. The dependent variable is the difference between the log of reported earnings and the income-elasticity-adjusted log of car values: $\ln E^{R} - [(1/\lambda)\ln C]$. Values are pooled ordinary least squares estimates, with standard errors clustered at the household (individual) level in parentheses. Private ownership and 1999 are the omitted categories. Ethnicity dummies are included in all regressions (coefficients are statistically insignificant and not shown). Excluded are car owners employed in federal and city government, law enforcement, military- and state-run education, and science and health care (households with at least one-third of total income from those sectors). The sample is restricted to car owners (households) with five or fewer cars who are 18 and older in 1999 and 60 and younger in 2003, excluding those earning less than the legal minimum wage or more than U.S.$100,000 in a given year and whose primary reported earnings in any year 1999–2003 came from nonwage sources. All specifications include industry dummies, but the coefficients are not shown. EED = employer earnings distribution.

$^{+}$ Significant at the 10% level.

* Significant at the 5% level.

** Significant at the 1% level.

dustries. The size of establishments varies across industries, and this variance, as already mentioned, is driven both by production technology considerations and by the choice to hide more. Since we cannot separate the production technology and hiding aspects in the optimal firm size choice, we elect to present estimations of regression (4) both controlling and not controlling for firm size. Coefficients on industry dummies in the specification not controlling for firm size are interpreted as capturing the relative composite effect of a particular industry, including the opportunities for hiding that are created by smaller firm sizes. Coefficients on industry dummies in the specification controlling for firm size, on the other hand, give us an idea of the relative cost of hiding across different industries that cannot be undone by the choice of the optimal firm size. We exclude foreign-owned firms from these estimations, and we also drop federal and Moscow city government sectors and public services (law enforcement, education, health care), which we analyze separately below. We also exclude mass media, private security firms, and the self-employed because the technology of hiding is very different in these occupations.

Table 8 presents the estimation results for selected industries of interest. To save space, we report only individual-level estimation results using our benchmark specification with $\lambda = .35$, but the results with $\lambda = .5$ are similar. Demographic controls and time fixed effects are included in all specifications, but the corresponding coefficients are similar to those in Table 7, so we omit them here.

The magnitudes of the coefficients on the log number of employees in Table 8 are very similar to those presented in Table 7, but the coefficients on many of the industry dummies are predictably sensitive to controlling for firm size. Looking first at the estimates in Table 8 that do not control for firm size, we see that trade and services in the private sector are estimated to report 60–70 percent less of their employees' earnings than other firms.[22] The cost of hiding in these industries is obviously reduced by the opportunity to conduct transactions using unregistered cash. Private banking hides about 40 percent more relative to other industries. State-owned utilities, transportation, and manufacturing, on the other hand, are estimated to report significantly more earnings for the same car values of their employees, which confirms our prediction that firms in these industries may find it very costly to engage in hiding behavior, although the magnitudes are somewhat different across household-level and individual-level data.

Once firm size is controlled for, the estimated coefficients decrease in magnitude. This is particularly pronounced in trade and services, as differences with other industries become statistically insignificant in the individual-level data. Apparently, the advantage of using unregistered cash in these industries is mostly related to the ability to choose a smaller firm size. The magnitude of hiding in

[22] We estimate the amount reported relative to the omitted category as $\exp(\beta_k)$, where β_k is the coefficient on the industry k dummy in Table 8.

Table 8
Estimates of Regression (4): Some Major Industries

	Household Level, λ = .74		Individual Level, λ = .35	
	(1)	(2)	(1)	(2)
Log number of employees		.242**		.361**
		(.015)		(.028)
Private sector:				
Banking and finance	−.464**	−.245+	−.435+	−.338
	(.139)	(.136)	(.261)	(.257)
Construction	−.252*	.146	−.249	.154
	(.122)	(.117)	(.230)	(.226)
Wholesale and retail trade	−1.239**	−.499**	−1.145**	−.292
	(.117)	(.120)	(.219)	(.225)
Manufacturing	−.520**	−.133	−.450*	−.013
	(.111)	(.108)	(.222)	(.220)
Services	−1.063**	−.341**	−.946**	−.100
	(.125)	(.125)	(.230)	(.235)
Communications and IT	−.498**	.103	−.412	.171
	(.180)	(.172)	(.288)	(.283)
State owned:				
Construction	−.112	.246	.019	.225
	(.194)	(.189)	(.314)	(.316)
Utilities	.470**	−.165	1.179**	−.010
	(.177)	(.184)	(.308)	(.327)
Transportation	.427**	.155	1.192**	.505+
	(.164)	(.163)	(.277)	(.288)
Manufacturing	.246	.190	1.566**	1.267**
	(.215)	(.217)	(.416)	(.421)
Constant	−6.030**	−7.425**	−21.062**	−22.713**
	(.299)	(.303)	(.450)	(.463)
N	11,017	11,017	11,927	11,927
Adjusted R^2	.108	.161	.080	.116

Note. The dependent variable is the difference between the log of reported earnings and the income-elasticity-adjusted log of car values: $\ln E^R - [(1/\lambda)\ln C]$. Values are pooled ordinary least squares estimates, with standard errors clustered at household (individual) level in parentheses. All regressions include age, gender, ethnicity, the number of members in the household, the number of earners in the household, and year dummies. The sample is restricted to car owners employed by private Russian firms (households with at least two-thirds of household earnings coming from the private sector) with five or fewer cars who are 18 and older in 1999 and 60 and younger in 2003, excluding those earning less than the annual equivalent of the legal minimum wage or more than the equivalent of U.S.$100,000 in a given year and those whose primary reported earnings in any year 1999–2003 came from nonwage sources. The omitted category is all other industries.

+ Significant at the 10% level.
* Significant at the 5% level.
** Significant at the 1% level.

banking also decreases somewhat when size is controlled for, but it does not decrease nearly as much as in other industries. This is consistent with the banking and finance industry having access to low-cost hiding technology that uses loopholes in the financial system and thus depends relatively less on being smaller.

Overall, these results seem to be consistent with our prior results about differential costs of hiding across different sectors of economic activity. These estimates, however, allow us to calculate income hiding relative only to some omitted sector. To measure the absolute magnitude of income hiding, in Section 4.3 we measure it relative to a subsample in which incomes are reported truthfully.

4.3. *Income Hiding and Corruption Relative to Foreign-Owned Firms*

We now use the group of Moscow car owners whom we assume to have more or less accurate earnings and who are employed by for-profit foreign-owned firms with 100 or more employees in each year as a benchmark to estimate the magnitudes of hidden earnings in different industries and the amounts of side incomes in government employment and public services. Since we do not have household-level data for the oversampled car owners employed by foreign-owned firms, we present the estimation results based on only the individual-level data with $\lambda = .35$ (see Section 5.1 for alternative estimations using household-level data and NLSY79 survey data as a benchmark, which produce similar magnitudes of hidden earnings). Estimates using as the benchmark a subsample of car owners with relatively higher car values (such as those employed in foreign-owned firms) are likely to result in an upward bias in coefficient β if the true elasticity of demand is actually higher than the one assumed in estimations, thus resulting in underestimation of actual income hiding. Consistent with this, we obtain even larger estimates of hiding in almost all industries when we perform estimations using the value of $\lambda = .5$ implied by the NLSY79 data.

4.3.1. How Much Do Different Industries Hide?

We interact private and state ownership dummies with industry (sector of employment) variables to distinguish between the amounts underreported in private and state-owned firms within the same industry. Estimated coefficients in regression (4) on all state-industry dummies for all years and for each year are reported in Tables A6.1–A6.4 of the online appendix. To make the magnitude of hidden earnings easier to assess, in Table 9 we convert estimated coefficients (for all years) into implied fractions of reported earnings, and we indicate coefficients for hidden earnings that are statistically significant at least at the 5 percent level.[23]

In the majority of industries in both the private and state-owned sectors, estimated fractions of reported earnings among car owners are very low. For example, private banks are estimated to report only 12.2 percent of employees' earnings, while private businesses in higher education and research, health care, mass media, retail and wholesale trade, and services all report less than 10 percent. There are exceptions, however. State-owned manufacturing companies and private secondary and earlier educational establishments are estimated to have reported earnings that are statistically indistinguishable from the benchmark,

[23] The number for industry k in Table 9 is equal to $\exp(\beta_k)$, where β_k is the coefficient on the industry k dummy in Table A6.1 in the online appendix.

Table 9

Fraction of Reported Earnings Relative to the Benchmark: 1999–2003

	Private		State	
	N	Fraction	*N*	Fraction
Banking and finance	1,021	.122*	126	.292*
Federal government	N.A.	N.A.	475	N.A.
City government	N.A.	N.A.	244	.374*
Law enforcement	N.A.	N.A.	325	.258*
Higher education and research	298	.069*	1,099	.176*
Secondary and earlier education	(35	.886	305	.163*
Health care	132	.094*	635	.239*
Mass media	279	.093*	103	.323
Construction	1,767	.151*	323	.200*
Utilities	41	.239	420	.654
Transportation and warehousing	353	.293*	444	.645
Wholesale and retail trade	2,215	.059*	137	.123*
Manufacturing	1,968	.122*	299	.937
Sports and entertainment	107	.071*	111	.130*
Services	1,746	.074*	149	.317*
Communications and IT	489	.124*	160	.179*
Self-employed	37	.049*	N.A.	N.A.
Noneducation not-for-profit	217	.102*	24	.124

Note. Data for all years are pooled and regression coefficients converted to implied fractions of reported earnings. The benchmark consists of car owners employed by a subsample of foreign-owned employers. The underlying regressions include age, gender, percentile in employer earnings distribution and its squared term, and the constant term. The sample is restricted to car owners (households) with five or fewer cars who are 18 and older in 1999 and 60 and younger in 2003, excluding those earning less than the legal minimum wage or more than U.S.$100,000 in a given year and those whose primary reported earnings in any year 1999–2003 came from nonwage sources. N.A. = not applicable.

* Estimate for the fraction of hidden earnings is significant at least at the 5% level.

and state-owned enterprises in utilities and in transportation and warehousing have fractions of reported earnings of about .65, but we cannot reject the hypothesis that they are actually the same as the benchmark.

Further examination of these fractions by year reveals some interesting patterns over time (see Tables A6.2–A6.4 in the online appendix for details). There is an overall trend toward more transparent income reporting. However, a closer look at the dynamics of these fractions reveals that the progress was very uneven and was mostly confined to the industries identified in the multiyear data as relatively more transparent: utilities, manufacturing (especially state owned), and state-owned transportation and warehousing. There is almost no change for sectors such as private retail and wholesale trade or services, while in private banking, after some improvement in 2001 (the year income tax reform was implemented), income reporting once again reverts to indicate more hiding.

As mentioned, in this particular case, we were able to compare our estimates on actual data for the year 2003 with the help of a friend with insider knowledge. The evidence he provided to us (see Table 1) shows an average fraction of reported earnings of just over 10 percent, slightly less than our estimate of 13 percent for the Moscow banking sector as a whole.

4.3.2. Side Earnings in Government and Public Services

Black wages in envelopes are unlikely to play a major role in hidden earnings of car owners employed by the federal and Moscow city governments and law enforcement. Unreported earnings in these sectors are likely to come from unregistered and illegal income sources, such as bribes and corruption. For public servants, such as educators and health care specialists, unregistered earnings represent more of a gray area. Those incomes could include more or less legitimate unreported side earnings, such as those from tutoring. Still, there is ample reason to believe that a lot of these hidden earnings come from illicit fees charged for public services that should be provided for free.[24]

As can be seen from Table 9, the biggest fraction of hidden earnings as compared to the benchmark (83.7 percent) is found in public secondary and presecondary education. This does not necessarily come as a surprise. Even though law enforcement tops the list of institutions perceived as "extremely corrupt" in Russia, with 49 percent of respondents subscribing to this view, according to the Global Corruption Barometer,[25] the education system also has a very high score, with 27 percent of respondents describing it as "extremely corrupt." And while the number of bribes collected by traffic police may be large, the amount of payment per bribe is not that high. In contrast, parents desperate to get their children into prestigious schools are ready to pay large amounts of money to corrupt educators.

Estimated coefficients over the years between 1999 amd 2003 (in the online appendix) also show that some sectors of public services experienced a noticeable decline in the fraction of unreported earnings among car owners over time. This is characteristic of the federal government, where we estimate that in 1999 about 80 percent of earnings came from unreported sources, but by 2003 earnings became statistically indistinguishable from the benchmark. The same trend, although to a lesser extent, can be observed in law enforcement and secondary education.

It is well known that in the early 2000s, during the first few years of the first term of Vladimir Putin, Russian authorities made attempts to fight corruption pervading the public sector. In particular, federal government employees and educators received substantial pay increases, which were coupled with tougher enforcement of legal and regulatory norms. Our data suggest that these measures did have an effect, at least at the time. We also detect almost no improvement in transparency among the car owners employed by the Moscow city government,

[24] For example, as this paper was being written, a scandal erupted in one of the elite Moscow secondary schools, where the director and his colleagues were suspected of extorting over $7 million from parents over the course of 3 years, collecting illegal fees for everything from admission to school to extracurricular activities (see Lenta, Director of Moscow School Suspected in Large-Scale Fraud [in Russian] [http://lenta.ru/news/2010/09/30/school]).

[25] See Transparency International, Global Corruption Barometer 2005, VOP TI REG STANDARD.xls (http://archive.transparency.org/content/download/5084/29852/file/gcb2005_tables_results_05_11_29.zip).

which remained insulated from the federal authorities under the charismatic Mayor Yuri Luzhkov.[26]

4.4. External Validity of the Sample of Car Owners

In our analysis above we document considerable income hiding in private domestic firms and parts of the government sector. Our identification strategy is based on using car values as a proxy for actual incomes. This naturally restricts our analysis to the sample of car owners. As we saw above, car owners tend to be positioned higher in an employer's hierarchy and thus represent an elite part of the workforce. Hence, determinants of transparency for this subsample of workers might be important on its own. Still the question remains to what extent the results obtained for this sample might be applicable to the general workforce in Moscow and beyond.

One immediately obvious dimension along which car owners are different from those who do not own cars is the revealed preference for owning a car. There is little reason to believe that such preferences would be systematically related to the nature of labor contracts offered to workers by their employers, and there is even less reason to believe that any such differences are related to labor contracts in systematically different ways between domestic and foreign-owned employers. Otherwise, any possible differences based on car preferences would cancel out when we compute income hiding by comparing car owners employed by foreign-owned companies with those employed by domestic firms.

To examine how much our findings may apply beyond the population of car owners, we use an individual's percentile (ranking) in his or her employer earnings distribution (EED) to obtain a proxy for that worker's position in the firm's hierarchy. We saw (Section 3.3.1; see also Section 2) that car owners tend to be higher paid employees in a company, as is indicated by their higher position in their EEDs. Does this imply that such employees are less transparent? To answer this question, we looked at the variation in transparency of car owners as a function of their relative position within the firm.

In Table 7 we saw that the percentile in EED has a positive but nonlinear effect on transparency. To analyze these effects less parametrically, we reestimate equation (4) by dividing the percentile range observed among car owners into 10 bins, assuming constant effect within each bin. We use simple ordinary least squares (OLS) regression and include employer fixed effects to absorb firm-level heterogeneity. To mitigate the potential impact of idiosyncratic shocks to current income, which would tend to raise an individual's position in the EED, we present the results for previous-year percentiles.[27] We also conduct an analysis using 5-year mean percentile dummies and find similar patterns. In particular,

[26] Luzhkov was finally removed from office in 2010 by then- president Dmitry Medvedev, with the official reason stated as "loss of president's trust," a thinly veiled reference to rampant corruption in the city administration.

[27] Results for contemporaneous percentiles are similar but, as expected, larger in magnitude.

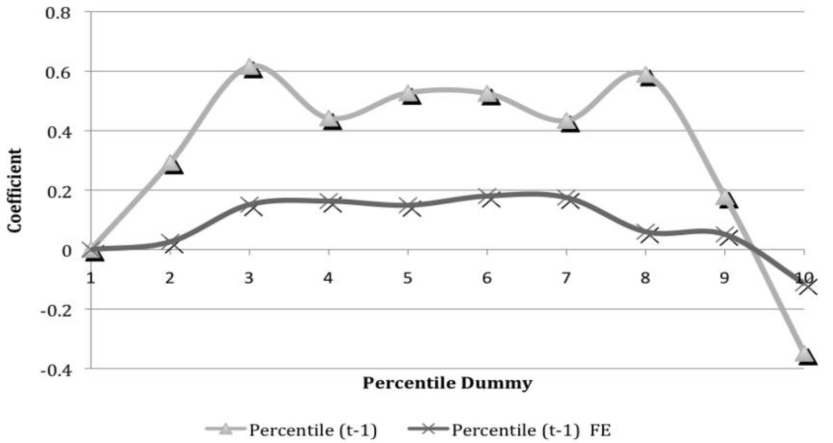

Figure 1. Income hiding and percentiles in employer earnings distribution: private sector

we estimate the following modification of equation (4):

$$\ln E^{\mathrm{R}}_{it} - (1/\lambda)\ln C_{it} = f_{j(i)} + \sum_{k=2}^{10} \beta_k P_{i,t}(k) + \gamma' \boldsymbol{X}_{it} + \varphi(t) + \varepsilon_{it},$$

where $P_{i,t}(k)$ is a set of dummy variables for whether a given individual's percentile in EED falls within a certain bin, $f_{j(i)}$ is employer fixed effects, and $\boldsymbol{X}_{it}$ includes age, age squared, a gender dummy, year fixed effects, and (log of) firm size.

Figures 1 and 2 plot estimated effects for each of 10 percentile bins for private companies and state-owned firms (estimated coefficients are presented in Section A11 of the online appendix).[28] For private domestic companies (Figure 1), the effect of percentile in the EED on transparency is increasing over the first three deciles and remains essentially flat from the fourth to eighth decile before decreasing at the very top of the EED. The patterns in SOEs and government services are similar (see Figure 2), but the picture is noisier, which is not surprising given that much of the hidden earnings in this sector comprise idiosyncratic incomes from corruption. We can thus detect a tendency for lowest paid employees and top management to be less transparent than employees in the middle, but overall, as we move from lower to higher paid employees in the same company, estimated transparency changes little for most of the range. Thus, there is no evidence that an average car owner higher in the EED would be less

[28] In Figures 1 and 2, coefficients are based on estimates of the relation between the difference between the log of reported earnings and the income-elasticity-adjusted log of car values, $\ln E - {}^{\mathrm{R}}[(1/\lambda)\ln C]$, and dummies for (lagged) percentiles in employer earnings distribution.

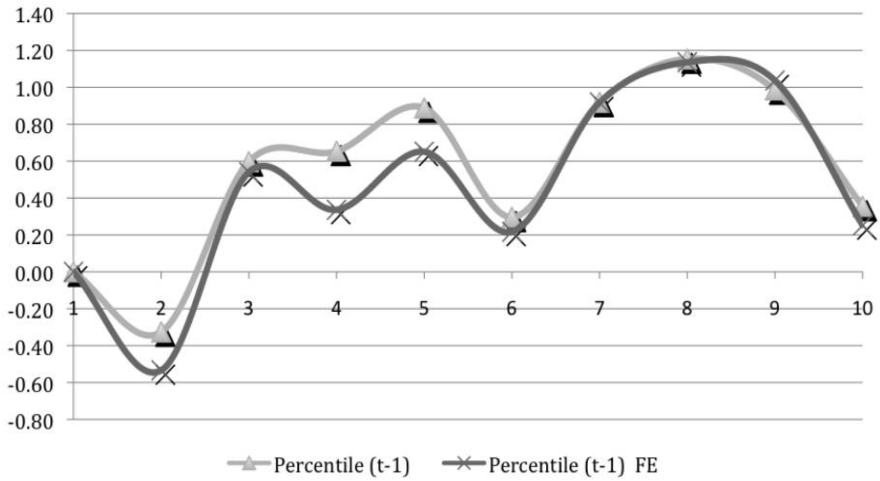

Figure 2. Income hiding and percentiles in employer earnings distribution: state-owned sector.

transparent in the way his or her income is reported by the employer than an average employee who does not own a car and is lower in the EED.[29]

5. Extensions and Robustness

5.1. Hiding Relative to an Alternative Benchmark: Household-Level Data Compared with the National Longitudinal Survey of Youth Data

In the benchmark sample of car owners employed by foreign-owned firms, we do not have data on household members. To see if taking into account household-level car ownership changes our estimates, we estimated regression (4) using household-level data on our main sample, with the NLSY79 data on U.S. households as the benchmark. This also serves as a strong robustness check of our earlier estimation results, as a benchmark is totally different in nature

[29] The decrease in transparency at the very top suggests that including top management might bias our results toward overestimating hidden earnings for the general workforce. We therefore reestimated, dropping the top 5 and top 10 percent of employees in the within-employer earnings distribution, and obtained very similar estimates. Details are available on request. Another limitation of our analysis is that it is based only on the data from Moscow. Even if Moscow is not representative of the entire country, it is the major political and economic center of Russia, officially accounting for more than 20 percent of total (reported) gross domestic product (GDP). We compared our findings with Mironov (2013), a recent study of tunneling and corporate theft, and found a high level of consistency between our estimates of hidden earnings and his estimates of tunneling and corporate fraud.

from the benchmark composed of employees of foreign-owned firms in Moscow used in Section 4.

To make the two data sets as comparable as possible, we limit the sample to U.S. households in central cities in their MSAs. We include only those employment sectors in Moscow and the NLSY that can be completely mapped onto each other: banking and finance; construction; utilities, transportation, and warehousing; wholesale and retail trade; and manufacturing. Finally, we include only Moscow households with wage earners in the same age bracket (38–46 years old in 2003) as in the NLSY79 data.

We estimate equation (4) using the income elasticity of demand $\lambda = .74$ obtained from the NLSY79 household-level data, as in Section 3. We interact industry dummies with private- and government-ownership dummies in the Moscow data, as in Table 9. The coefficients on ownership-industry dummies and the implied fractions of reported earnings in the Moscow data are reported in Table 10.

Despite the fact that the estimates in Table 10 use a very different benchmark and rely on household- rather than individual-level data, the similarities between the estimates in Tables 9 and 10 are very strong. Estimation results with the NLSY data as the benchmark have very similar or even lower fractions of reported earnings in all ownership-industry categories in the Moscow household-level data, with the exception of state-owned trade establishments. The weighted average fraction of reported earnings is also similar between the two estimates: 14.7 percent with the NLSY79 data as the benchmark and 18.9 percent with Russians employed by foreign-owned firms as the benchmark.

5.2. Sensitivity to Estimation of Car Demand

Our identification strategy relies in part on having the right estimate for the income elasticity of the demand for the stock of cars, λ. Hence, it is important to examine how sensitive the results are to possible errors in this estimation and to a more complicated relationship between (true) earnings and car values, including possible threshold effects. It is also possible that the value of λ varies across individuals and households, so it is important to know how much our results might be affected if we allow such heterogeneity.

For the sake of parsimony, we relegate the detailed description of how we address this issue to the online appendix and briefly discuss the tests and results here. In Section A7 of the online appendix we formally examine the direction and the boundaries of a possible bias that may be introduced into our estimates of income hiding by setting λ at a value lower than its true value λ^* ($\lambda < \lambda^*$) (the opposite case can be derived symmetrically). In all cases in which car owners in the benchmark sector, foreign-owned firms, own more expensive cars than car owners in our industry or sector of interest, our estimate of income hiding will be biased downward if we assume the income elasticity of the demand parameter to be less than it actually is. We also reestimated our main regression

Table 10

Fraction of Reported Earnings in Moscow Household-Level Data, with National Longitudinal Survey of Youth as the Benchmark

	Private	State
Banking and finance:		
Coefficient	−2.086**	−1.138**
SE	.180	.394
Fraction	.124*	.321*
N	318	50
Construction:		
Coefficient	−1.945**	−1.837**
SE	.147	.213
Fraction	.143*	.159*
N	672	156
Utilities, transportation, and warehousing:		
Coefficient	−1.919**	−1.574**
SE	.241	.188
Fraction	.147*	.207*
N	176	364
Wholesale and retail trade:		
Coefficient	−2.784**	−1.677**
SE	.154	.347
Fraction	.062*	.187*
N	661	56
Manufacturing:		
Coefficient	−2.270**	−1.500**
SE	.158	.274
Fraction	.103*	.223*
N	714	152

Note. Fractions of reported earnings are calculated from the corresponding coefficients on ownership-industry dummies. The benchmark (omitted category) consists of car-owning respondents to the National Longitudinal Survey of Youth 1979 cohort 2000 and 2004 surveys, limited to those residing in central cities in metropolitan statistical areas and employed in the following five industries: banking and finance; construction; utilities, transportation, and warehousing; wholesale and retail trade; and manufacturing. Regressions include age, gender, number of members in the household, year dummies, and the constant term. The Moscow sample is restricted to households with five or fewer cars and earners between 38 and 46 years old in 2003, excluding those earning less than the legal minimum wage or more than U.S.$100,000 in a given year and whose primary reported earnings in any year 1999–2003 came from nonwage sources. The total number of observations is 4,215, and the adjusted R^2-value is .209.

* Estimate for the fractions of hidden earnings is significant at least at the 5% level.

** Significant at the 1% level.

equation (4) using a semiparametric approach in which we let the parameter λ take different values for different income brackets, allowing the elasticity of the demand for cars to vary with income, and found very similar results.

We also conducted tests to address the issue of possible unobserved heterogeneity in individual demand for cars and the demand for cars by car owners employed in different sectors (for example, because of differences in technological need for cars or show-off culture across sectors). None of our findings are seriously affected by taking these factors into account.

5.3. Ownership of Prestigious Cars

Owning a prestigious car was and remains one of the most sought-after symbols of high-income status in Russia, so examining the distribution of ownership of such cars could reveal a lot about real income differentials while serving as another powerful robustness check that does not depend on our assignment of car prices.[30] In Table 11 we show the fraction of prestigious and nonprestigious cars owned by employees in five employer-size quintiles in the private sector and some select sectors of employment, along with their corresponding mean reported earnings.

As can be seen from Table 11, as with car values, the relationship between earnings and the fractions of prestigious and nonprestigious cars is reversed for firms of small and large size. Employees of private firms in the smallest size quintiles own a fraction of prestigious cars that is more than double that of employees of private firms in the largest quintile, but their reported earnings are less than half those of owners of nonprestigious cars employed in large firms. Reported earnings are substantially higher for owners of prestigious cars than for owners of nonprestigious cars within each firm-size quintile but not necessarily across quintiles. This renders support to our hypothesis that each employer reports a fraction of earnings paid to its employees (so that the ordering of reported and true earnings is preserved within each employer), but these fractions vary systematically across employer types.

Turning to sectors of economic activity, we find again that employees in law enforcement and public secondary education and health care have fractions of prestigious cars that are higher than those of federal government employees, even though the former sector's reported earnings are just a fraction of the latter's. Car owners employed in wholesale and retail trade have a much higher fraction of prestigious cars than that for communications and IT, but their reported earnings are much lower, and so on. It is interesting to note the reversal of relative reported earnings between owners of prestigious and nonprestigious cars in law enforcement and public secondary education and health care. This can be interpreted as additional evidence of the importance of corruption (bribes) as the source of unreported income in those sectors rather than systematic employer-level underreporting in the private sector.

To illustrate income-hiding patterns even more clearly, we also consider a sample of individuals who own cars worth $25,000 or more in any given year and who stayed for 5 years in the same ownership-industry sector of employment. For car owners employed in the Russian private sector, the average reported earnings were $7,509, while the average car values were $43,222; for the state-

[30] We assign the following vehicles to the prestigious category: higher end models of Audi, BMW, Mercedes, Volvo, Saab, Lexus, and Infinity; some high-end models of Alfa Romeo, Rover, Renault, Toyota, Ford, General Motors, and Chrysler; and all models of Cadillac, Porsche, Jaguar, and Bentley. Nonprestigious cars are Russian-made cars and Chinese and Korean cars (the status of the latter two was still quite low in Russia at the time of our analysis). The full list is available on request.

Table 11
Prestigious and Nonprestigious Cars and Earnings, by Employer Characteristics

	Fraction		Owners' Annual Earnings	
	Nonprestigious Cars	Prestigious Cars	Nonprestigious Cars	Prestigious Cars
Firm size quintile (private sector):				
First	53.65	18.71	1,413	2,335
Second	62.34	15.37	2,301	4,150
Third	67.09	12.80	3,000	6,086
Fourth	71.46	9.93	4,047	7,455
Fifth	74.37	8.98	4,934	9,158
Sector of employment:				
Federal government	71.32	9.04	5,156	5,289
Law enforcement	78.06	11.22	2,457	2,426
Public secondary education and health care	71.25	11.19	2,419	1,803
Private trade and services	60.78	15.14	2,341	3,355
Private communications and IT	58.49	11.81	5,314	8,816

Note. Nonprestigious cars are Russian-, Korean-, and Chinese-made cars. Prestigious cars include luxury and exotic brands (Porsche, Bentley, Jaguar, and so forth); higher end models of Audi, BMW, Mercedes, Volvo, Saab, Lexus, and Infinity; and some high-end models of Alfa Romeo, Rover, Renault, Toyota, Ford, General Motors, and Chrysler. Annual earnings are in U.S. dollars. IT = information technology.

owned and public services sector, the corresponding values were $6,214 and $41,919, while for the foreign-owned sector they were $32,334 and $35,377. Thus, even among this select sample of owners of expensive cars, reported incomes were four times higher in the foreign-owned sector than in the Russian private sector (and six times higher than in the state-owned and public services sectors), while the car values were 20–25 percent lower. Furthermore, in the Russian private sector, the car values for owners of the most expensive cars were almost exactly the same among those employed in the smallest size quintile and the largest size quintile ($44,509 and $44,760, respectively), but their reported incomes differed by a factor of five ($3,657 in the smallest size quintile versus $17,957 in the largest size quintile). Hence, if their reported earnings are to be believed, car owners in the smallest Russian firms had to save all of their reported earnings for more than 12 years to be able to afford their cars. (The results are very similar for other sectors, where we found especially large fractions of hiding in our estimations in Section 4.)

5.4. Other Robustness Checks

Our empirical specification in equation (4) and the estimations presented previously did not include individual or household fixed effects. The reason is that given the short duration of our panel and stickiness of the demand for cars, fixed effects may produce biased estimates of the coefficients on our variables of interest. Nevertheless, we did check how our results are affected by controlling for individual fixed effects. These results, with foreign-owned companies as a benchmark, are presented in Section A9 of the online appendix. As expected, the implied fractions of reported earnings are higher when fixed-effect estimates are used. For example, employment in the private wholesale and retail trade, which is associated with hiding 94 percent of earnings in our preferred pooled OLS estimation, is associated with hiding only 62 percent of earnings in fixed-effect estimation. Similarly, the estimated fraction of reported earnings in private banking using fixed-effect estimation is 69 percent as opposed to 88 percent in pooled OLS estimates, in private services it is 73 percent as opposed to 93 percent, and so on. We have also reestimated the regressions in Section 4 using 5-year averages of incomes as a proxy for permanent income, while limiting the samples to workers who did not cross ownership and sector lines of employment over those 5 years. The estimation results were once again very similar.

6. Implications for Official Statistics of Gross Domestic Product

Before its collapse in 1991, the Soviet Union had been viewed as one of the world's two superpowers. According to the Penn World Table (PWT) (Heston, Summers, and Aten 2012), in 1990 Russia's GDP per capita was still estimated to be about 46 percent of the U.S. level. But according to the same source, in 1999 the estimated Russian GDP per capita was only about 23 percent of the U.S. level, which implies that in less than 10 years it had lost half of its economic potential relative to the United States.

In hindsight, the strength of the Soviet economy had perhaps indeed been exaggerated. But our estimates, based on car ownership, suggest that the analysts may have gone too far in the opposite direction during the first decade of the Russian transition to a market economy.[31] Official sources have put the fraction of the shadow economy in Russia at about 25 percent,[32] and while some economists have come up with larger numbers (for example, 45.6 percent in Alexeev and Pyle [2003], using electricity consumption data), our estimates indicate that it might be even much higher than that. Even though our sample of car owners in Moscow may not be representative of the whole Russian economy, we can at least reassess the economic role of Moscow in the country's economy. In official statistics, the share of Moscow in Russian GDP was 18.2 percent in 1999, in-

[31] Young (2012) uses data on ownership of durables to estimate actual growth rates in sub-Saharan Africa and also comes up with considerable upward revisions compared with Penn World Table (PWT) estimates.

[32] In a development little known among nonspecialists, the State Committee of the Russian Federation on Statistics (Goskomstat) since the mid-1990s has been adding about 25 percent to the data it collects to correct the GDP estimates for the presence of the hidden economy. While the practice seems to date back to at least 1995, the rather elaborate methodology employed was formalized in 1998 (see Federal State Statistics Service, On the Adoption of Basic Methodological Principles of Assessing the Hidden (Informal) Economy [in Russian], Ordinance No. 7 [January 31] [http://www.lawru.info/legal2/se5/pravo5464/]).

creasing to 20.4 percent in 2003.[33] But if Moscow employers hide 80 percent of the incomes they generate, this could mean that Moscow's true share of Russia's GDP could be anywhere up to 55 percent in those years (depending on how much hiding there is in the rest of the country).

To get a general idea of what our estimates of hiding imply for actual GDP per capita in Russia relative to the United States, let us assume, following Alexeev and Pyle (2003), that the fraction of hiding in the rest of the country is indeed 45.6 percent. A simple back-of-the-envelope calculation, then, results in an estimate of Russian GDP per capita relative to the United States in 1999 of 41.6 percent, rather than 23 percent as in PWT.[34] Comparing this number to the PWT estimate for 1990, we can see that instead of a decline in living standards of one-half relative to the United States during the first decade of the transition to a market economy, Russia's per capita GDP relative to the United States in fact declined very little (if at all) over the same period. This makes more sense from an intuitive perspective, given that a lot of the decline in output observed between 1990 and 1999 occurred in military- and other state-procurement-related sectors (including intermediate inputs, a huge waste for which the Soviet economy was infamous), while output of consumption goods and especially services, including financial and other services that had not existed prior to the collapse of communism, had greatly expanded.

Our findings also cause us to reassess the degree of resurgence of the Russian economy in more recent years (at least until the worldwide financial crisis and the collapse in primary resources prices in late 2008). The officially measured real GDP index had returned to 70.7 percent of its 1990 level by 2001, to almost 80 percent by 2003, and to 96.7 percent by 2006. The nominal GDP converted to U.S. dollars at the market exchange rate increased by 152 percent between 1999 and 2003 for the country as a whole and by 188 percent for the city of Moscow.[35] Officially reported earnings in our data record very similar increases over the same period (161 percent in U.S. dollars on average for private firms and 184 percent on average for the government and state-owned sectors). The implied growth in car values (with income elasticity equal to .35, as in our

[33] See Federal State Statistics Service, Gross Regional Product [in Russian] (http://www.gks.ru/free_doc/new_site/vvp/vrp98-12.xls).

[34] As already mentioned, the PWT (Heston, Summers, and Aten 2012) estimates Russian GDP per capita to be 23 percent of the U.S. level in 1999, but we subtract 25 percent because of the correction for the hidden economy already applied by the State Committee of the Russian Federation on Statistics. Excluding this adjustment, Russian GDP per capita would thus be $.75 \times 23 = 17.3$ percent of the U.S. level. To obtain 41.6 percent, we then multiplied 17.3 by the sum of (i) the 18.2 percent share of Moscow in official statistics, divided by .2—the fraction of reported earnings we estimate here—and (ii) the 71.8 percent share of the rest of the country in official statistics, divided by .544—the fraction of reported earnings estimated by Alexeev and Pyle (2003), applied to the rest of the country. Note that the correct share of Moscow in Russian GDP in 1999 under these assumptions would be 37.7 percent, a number that also has more intuitive appeal than the official 18.2 percent.

[35] Our calculations are based on official statistical data from Federal State Statistics Service, Gross Regional Product [in Russian] (http://www.gks.ru/free_doc/new_site/vvp/vrp98-12.xls).

baseline specification) is 56.5 and 64.5 percent, respectively. Instead, car values grew by only 37 percent in the private sector and by 24.3 percent in the state-owned sector. These values suggest that increased transparency might indeed be responsible for a large share of growth in statistically measured GDP (compare Gorodnichenko, Martinez-Vazquez, and Peter 2009). And the increase in transparency seems to have been more pronounced in the state-owned than in the private sector.

7. Conclusions and Discussion

In this paper we propose a novel approach to measuring the hidden economy by juxtaposing reported individual incomes of car owners with their cars. Using administrative data on the earnings of Moscow car owners and data on their cars, we estimate that from 1999 to 2003 hidden earnings may have on average constituted 80 percent or more of the total wage earnings of car owners in a large number of firms in the private sector and in part of the government and public services sectors in Moscow. We also present some evidence that our results could be applicable more generally to the labor compensation of the Russian workforce and indicative of the larger phenomena of corporate theft and accounting fraud in Russian companies. These estimates have important implications for assessing the real size of the Russian economy and its actual standards of living, as well as the true allocation of resources across various sectors of economic activity.

Our results survive a variety of robustness checks. We consider individual- as well as household-level data, use alternative benchmark subsamples, look at different specifications of the demand for cars with constant and varying elasticity of demand, and use the panel nature of the data to perform estimation with fixed effects and time averages.

Our major contribution to the literature is estimates of hidden earnings disaggregated by employer type. We find that smaller firms hide more than large firms; private firms in trade, services, and banking hide more than SOEs in manufacturing and utilities; and foreign-owned firms hide dramatically less than domestic firms.

Our results on relative income hiding depending on observable employer characteristics are important in several respects. For example, one of our most robust findings is that when hidden earnings are taken into account, the firm-size effect on wages is reversed. Larger firms seemingly pay more to their workers, but in fact they pay less. In the industrial organization literature (that mostly uses theoretical models developed for institutionally sound market economies and the corresponding empirical evidence), larger firm size is associated with higher productivity. Our analysis indicates that although the official data from countries with less than optimal institutional environments may suggest the same, that is not necessarily the case in reality. At the same time, our study also sounds a note of caution with regard to policy measures promoting small businesses for development. Such measures, if not accompanied by provisions ensuring increased transparency and openness, can end up feeding the hidden economy.

Our findings have implications for the way in which the role of foreign direct investment is perceived in countries where, like Russia, nontransparency and corruption are rampant. According to the administrative data (which form the basis of official statistics), the performance of foreign-owned firms in Moscow was nothing short of miraculous as they managed to pay wages that were three to four times higher than those of domestic firms. But according to the corresponding car values of employees, the difference was much smaller, and the appearance of a miracle was due to stark differences in reporting.

We also found disturbingly large hidden earnings, most likely from corrupt sources, in all areas of government employment. A corrupt government and law enforcement not only lack credibility and moral authority to fight against tax evasion in the private economy; they also have incentives not to fight it seriously so as to increase the sources of their corrupt incomes. This problem will complicate Russia's push to a more transparent economy for years to come. But a potentially even more disturbing finding is that corruption seems to be at its worst in the state-run education system. In particular, concerned voices point out that the quality of education is deteriorating sharply. The large gap between the official and estimated earnings of educators that we found corroborates this view, albeit indirectly. The hidden economy is not just affecting Russia's present generation but also is corrupting its future generations.

Finally, our estimates indicate that official GDP statistics in Russia should be taken with a grain of salt. According to our rather conservative calculations, actual GDP per capita in Russia might have been considerably higher at the turn of the 21st century (up to twice as large) than official figures would suggest. Thus, both the decline in Russian GDP during the first decade after the collapse of communism under Boris Yeltsin and the subsequent recovery under Putin may have been seriously exaggerated by not taking into account changes in transparency of income reporting.

We conclude by noting some general lessons. Lacking firm statistical evidence, economists are understandably reluctant to embrace anecdotes suggesting that the hidden economy has a much larger scale than suggested by survey data or macroeconomic indicators. The increased availability of matched employer-employee data sets, and the relative ease with which data on personal assets such as cars are available and can be matched to those data sets mean that our methodology may find more widespread use. This can be expected to contribute to better understanding of the magnitude and, most important, the structural aspects of the hidden economy and corruption, which is badly needed for policy evaluations and recommendations.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. High Wage Workers and High Wage Firms. *Econometrica* 67:251–333.

Alexeev, Michael, and William Pyle. 2003. A Note on Measuring the Unofficial Economy in the Former Soviet Republics. *Economics of Transition* 11:153–75.

Andreoni, James, Brian Erard, and Johnathan Feinstein. 1998. Tax Compliance. *Journal of Economic Literature* 36:818–60.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie, and John Roberts. 2010. Why Do Firms in Developing Countries Have Low Productivity? *American Economic Review: Papers and Proceedings* 100:619–23.

Braguinsky, Serguey. 2009. Postcommunist Oligarchs in Russia: Quantitative Analysis. *Journal of Law and Economics* 52:307–49.

Braguinsky, Serguey, and Sergey Mityakov. Forthcoming. Foreign Corporations and the Culture of Transparency: Evidence from Russian Administrative Data. *Journal of Financial Economics.*

Brown, Charles, and James Medoff. 1989. The Employer Size–Wage Effect. *Journal of Political Economy* 97:1027–59.

Bush, Jason. 2004. More Russians Are Saying "Charge It." *Businessweek,* October 4. http://www.businessweek.com/magazine/content/04_40/b3902157_mz035.htm.

Cagan, Phillip. 1958. The Demand for Currency Relative to the Total Money Supply. *Journal of Political Economy* 66:303–28.

Ernst & Young. 2008. *Automotive Market in Russia and the CIS: Industry Overview.* Moscow: Ernst & Young. http://www.globalautoindustry.com/content/dl_presentation.php?dl_id=385&pid=15014.

Federal State Statistics Service. 2003. *Russian Statistical Yearbook 2003* [in Russian]. Moscow: Federal State Statistics Service. http://www.gks.ru/doc_2003/year03.zip.

———. 2005. *Labor and Employment in Russia, 2005.* Moscow: Federal State Statistics Service. http://www.gks.ru/bgd/regl/B05_36/Main.htm.

Feinstein, Jonathan S. 1999. Approaches for Estimating Noncompliance: Examples from Federal Taxation in the United States. *Economic Journal* 109:F360–F369.

Fiege, Edgar L., and Ivica Urban. 2008. Measuring Underground (Unobserved, Nonobserved, Unrecorded) Economies in Transition Countries: Can We Trust GDP? *Journal of Comparative Economics* 36:287–306.

Forbes Magazine (Russia). 2004. Bazu zdali [The base surrendered]. December 3.

Frey, Bruno, and Hannelore Weck-Hanneman. 1984. The Hidden Economy as an "Unobserved" Variable. *European Economic Review* 26:33–53.

Giles, David E. A. 1999. Measuring the Hidden Economy: Implications for Econometric Modelling. *Economic Journal* 109:F370–F389.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter. 2009. Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia. *Journal of Political Economy* 117:504–54.

Guriev, Sergei, and Andrey Rachinsky. 2008. The Evolution of Personal Wealth in the Former Soviet Union and Central and Eastern Europe. Pp. 134–49 in *Personal Wealth from a Global Perspective,* edited by James B. Davies. Oxford: Oxford University Press.

Harding, Luke. 2011. Abramovich v Berezovsky: What Have We Learned So Far? *Guardian,* November 7. http://www.guardian.co.uk/world/2011/nov/07/abramovich-berezovsky.

Heston, Alan, Robert Summers, and Bettina Aten. 2012. Penn World Table Version 7.1.

Philadelphia: University of Pennsylvania, Center for International Comparisons of Production, Income and Prices.

Holger, Görg, Eric Strobl, and Frank Walsh. 2007. Why Do Foreign-Owned Firms Pay More? *Review of World Economics* 143:464–82.

Hurst, Erik, Geng Li, and Benjamin Pugsley. 2012. Are Household Surveys like Tax Forms: Evidence from Income Underreporting of the Self-Employed. Working paper. University of Chicago, Booth School of Business, Chicago.

Ivanova, Anna, Michael Keen, and Alexander Klemm. 2005. The Russian "Flat Tax" Reform. *Economic Policy* 20:397–444.

Kaufmann, Daniel, and Aleksander Kaliberda. 1996. Integrating the Unofficial Economy into the Dynamics of Post-socialist Economies: A Framework of Analysis and Evidence. Policy Research Working Paper No. 1691. World Bank, Washington, D.C.

Klepper, Steven, and Daniel Nagin. 1989. The Anatomy of Tax Evasion. *Journal of Law, Economics, and Organization* 5:1–24.

Lyssiotou, Panayiota, Panos Pashardes, and Thanasis Stengos. 2004. Estimates of the Black Economy Based on Consumer Demand Approaches. *Economic Journal* 114:622–40.

Mironov, Maxim. 2013. Taxes, Theft, and Firm Performance. *Journal of Finance* 68:1441–72.

Mironov, Maxim, and Ekaterina Zhuravskaya. 2012. Corruption in Procurement and Shadow Campaign Financing: Evidence from Russia. Working paper. Instituto de Empresa Business School, Madrid.

Pissarides, Christopher A., and Guglielmo Weber. 1989. An Expenditure-Based Estimate of Britain's Black Economy. *Journal of Public Economics* 39:17–32.

Schneider, Freidrich, and Dominik Enste. 2000. Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature* 38:77–114.

Strobl, Eric, and Robert Thornton. 2002. Do Large Employers Pay More in Developing Countries? The Case of Five African Countries. Discussion Paper No. 660. Institute for the Study of Labor, Bonn.

Tanzi, Vittorio. 1983. The Underground Economy in the United States: Annual Estimates, 1930–1980. *IMF Staff Papers* 30:283–305.

Vedomosti. 2004. Dokhody v otkrytom dostupe [Earnings are in open access]. November 25.

———. 2005. Moskvichi pod kolpakom [Muskovites under surveillance]. May 26.

Young, Alwin. 2012. The African Growth Miracle. *Journal of Political Economy* 120:696–739.