

ECONOMIC
RESEARCH
FORUM



منتدى
البحوث
الاقتصادية

2008

working paper series

ESTIMATING THE SIZE OF TURKEY'S
INFORMAL SECTOR: AN EXPENDITURE
BASED APPROACH

Nurhan Davutyan

Working Paper No. 403

**Estimating The Size of Turkey's Informal Sector: An
Expenditure Based Approach**

Nurhan Davutyan

Working Paper 403

May 2008

Nurhan Davutyan: Dept. of Economics & Dept. of Industrial Engineering,
Marmara University
Email: dnurhan@eng.marmara.edu.tr, dnurhan@hotmail.com

Abstract

How extensive is the Turkish unofficial economy? An accurate answer to this question has considerable practical significance for various aspects of economic policy – particularly employment, fiscal and migration policies. Estimates based on macroeconomic data tend to be too imprecise. We utilize the most up to date household income and expenditure surveys to examine the extent of income underreporting among members of Turkey's unofficial economy. The Pissarides-Weber (1989) approach that we use, hypothesizes that survey data would reflect income underreporting in the informal sector as "excess food consumption." Since it relies on "marginal propensity to consume food", this method allows comparisons across time and space, via Engel's Law. Our results suggest that informal sector members spend more than their formal sector counterparts with the same level of reported income. Based on this information, we estimate the average size of the true informal sector income to be about 1.25 times the reported one. This implies the Turkish disposable income is $(25\%)*(83\%)$ or about 21% larger than the official estimate based on reported magnitudes.

ملخص

ما مدى اتساع الاقتصاد غير الرسمي التركي؟ الإجابة الدقيقة على هذا السؤال لها مغزى عملي ملموس بالنسبة لمختلف جوانب السياسة الاقتصادية في تركيا، لا سيما فيما يتعلق بالوظائف والضرائب والهجرة تبدو التقديرات التي تعتمد على بيانات الاقتصاد الكلي غير دقيقة. كما نستخدم أحدث الدراسات عن الدخل والإنفاق لدراسة التقصير في التبليغ عن الدخل بين الأفراد العاملين في الاقتصاد غير الرسمي في تركيا. ويفترض أسلوب بيساريدز - ويبر (1989) الذي نتبعه أن بيانات الاستقصاء من شأنها أن تعكس مدى التقصير في التبليغ عن الدخل في القطاع غير الرسمي باعتباره إفراطاً في استهلاك الطعام. ولأن هذه الطريقة تعتمد على استخدام طريقة "الميل الحدي لاستهلاك الطعام" نجدها تتيح المقارنات الزمانية والمكانية من خلال استخدام قانون إنجل. وتشير النتائج إلى أن أفراد القطاع غير الرسمي ينفقون أكثر من نظرائهم في القطاع الرسمي الذين يتلقون ذات الدخل. وبناء على هذه المعلومة فحن نقدر أن متوسط نسبة الدخل في القطاع غير الرسمي تصل إلى حوالي 1.25 من الدخل الرسمي ويدل هذا على أن الدخل المتاح في تركيا $(25\%)*(83\%)$ أو أكثر بنسبة 21% من التقديرات الرسمية القائمة على حجم الدخل المبلغ عنها.

1. Introduction

Turkey as an emerging economy has an informal sector, the magnitude of which is unknown. Based on common sense, casual observation and anecdotal evidence, there is a consensus that its size must be large (see the excellent survey by Ulgen and Ozturk (2006)). Essentially, a large pool of people seeking work – the result of an ongoing rural to urban migration – coupled with high taxes on employment gives rise to a large “unofficial or informal” sector comprising all sorts of legitimate activities which are beyond the reach of the authorities¹. Added to this is the inadequacy of the audit/supervisory structure. For instance whereas the number of tax auditors per 1000 inhabitant is 1.3 in France and 1.5 in England, the comparable figure for Turkey is 0.6. In addition the number of labor inspectors for the whole country of a population surpassing 70 million is (World Bank 2006). Thus, according to the same source, almost 53% of the workforce is outside the social security system.

This problem exists in many other countries. For instance after the fall of communism, efforts to develop market based tax systems in Eastern and Central Europe have faced difficulties of a similar nature. The total sum of income taxes, social security and compulsory pension contributions have resulted in high statutory marginal and average tax rates. When a certain point is reached², private sector firms and their employees collude to report a minimum salary and the rest of the compensation is paid in cash without any record. Understandably, such activity is more common in small and medium enterprises. In large companies whose shares are publicly traded, such arrangements are hard to implement. Besim and Jenkins (2005) provide further details. Another strategy is “home production” without any official reporting. Larger firms engage in “putting out/outsourcing relationships” with such micro enterprises. In Turkey they are known to be especially widespread in construction, textiles, catering and consulting, Ulgen and Ozturk (2006).

However, the overall damage to the system goes well beyond the personal income tax realm. Given the underreporting of labor costs, there will now be an overstatement of business taxable income. To correct this problem firms must now make unrecorded cash sales. This means domestic value-added taxes will also be underreported. Thus the end result of (very) high marginal and average tax rates on employees is to destroy/distort the business income and the domestic value added taxes in addition to the personal income tax system. This phenomenon is not restricted to Central and Eastern Europe. It exists in Latin America as well as in Turkey and other Middle Eastern-North African countries. For a macro oriented discussion, see Henry and Springborg (2001, pp75-87) and Djankov et al (2002) for a detailed and micro oriented overview. Al-Kawaz (1999) surveys the informal sector in a number of Arab countries. From a national income accounting perspective, the net result is a downward biased GDP figure.

2. Causes of Income Underreporting in Turkey.

High overall employment taxes, coupled with a low enforcement capacity – both a low probability of detection and the inadequacy of penalties if detected – are the chief reasons for firms to resort to unreported/unofficial activities. The World Bank’s already cited Labor Market Study documents the weaknesses of the relevant supervisory structure. When it comes to employment taxes themselves, according to OECD, Turkey has the highest tax burden.

¹ Seyfettin Gursel: “Interview with Economics Minister Ali Babacan” in daily VATAN 01/12/2006.

² In microeconomic terms, when the marginal factor cost (inclusive of such taxes) of labor exceeds its marginal product value, this point is reached.

As can be seen from Figure 1, the ratio of employment taxes to overall employee compensation in Turkey is 43%. This compares with an OECD average of 26%. To understand “the fatal attraction” of informality for employers, one should add to this direct burden non-tax financial obligations such as social security contributions, health insurance and environmental requirements. In the presence of a large pool of unemployed workers, weak legal enforcement and competitive pressures from low-wage countries, many firms turn to unofficial production.

Another issue that has to be confronted involves the trade-off between employment security and employment level. In other words laws and regulations aimed at providing job security to existing workers may lead to labor market inflexibilities and thus harm the economy’s job creation capacity. To put it more bluntly, faced with high firing costs employers may prefer to hire less in general or to hire less in the formal sector and more in the informal sector. From this perspective Sapir (2005) distinguishes among two types of employment policies in the EU: Nordic vs. Mediterranean. The first Scandinavian model with low job security but high and flexible unemployment support gives the best results in terms of both performance and employment levels. The second one, practiced in countries like Spain, Portugal or Greece, provides high job security but low unemployment support. He argues that this policy leads to low productivity and low employment.

As seen from Table 1, in terms of hiring and firing costs, Turkey resembles the Mediterranean countries. For instance, next to Portugal it has the highest firing costs. But unlike them Turkey is still undergoing urbanization, Derviş et al (2004). The presence of a large pool of unskilled or semi-skilled rural migrants, combined with high employment costs is giving rise to an unofficial economy. In addition Turkey’s high firing costs hamper job creation. Faced with such high costs, employers prefer overtime work to new hiring. As a result, in 2006 the Turkish industry with 52 hrs/week had the highest weekly working hours among OECD members. The comparable EU-15 average was 38.5 hours and that for Korea was 48 hours. In other words if the Korean figure were to prevail in Turkey, there would have been 325,000 more jobs, World Bank (2006).

The compulsory minimum wage is another factor feeding informal employment. Figure 2 shows that minimum as a ratio to per capita GDP. Turkey had the highest relative minimum wage compared to a wide variety of countries.

Another factor contributing to the rise of the informal sector involves bureaucratic hurdles confronting new business creation. For Turkey, licensing procedures seem to need improvement. Using figures from World Bank’s Doing Business Database (2006) the average number of licensing procedures in low-average income countries is 18, whereas the comparable figure for Turkey is 32.

Finally social consciousness in Turkey does not view informality and tax evasion as reprehensible. This is due to a lack of transparency in public administration and the concomitant lack of trust between state and civil society. According to the results of a field study by Adaman et al (2001) “In Turkey the population’s trust in public institutions is very low. There is widespread and high level of dissatisfaction about the level and quality of public services...More than half of the interviewees complained about the services provided by the traffic police, tax/fiscal personnel and customs officials.” Such attitudes imply a psychological climate propitious to unofficial economic activity and the resulting tax evasion.

3. Previous Research

In their survey on measuring “the unofficial economy” Schneider and Enste (2000), discuss three broad measurement categories: direct, indirect and model-based approaches. The first approach utilizes direct evidence in the form of surveys or tax audits to estimate the size of the informal sector. We do not know of any study applying this method to Turkish data.

The “indirect” methods use various economic indicators – such as the money stock, employment level – or key physical indicators – such as electricity – to reach a conclusion about the magnitude of the unofficial economy. Note that these are all macroeconomic aggregates. As such they necessitate the use of time series data. There are various studies applying one or another form of the “indirect” methods to Turkey. Due to data imperfections and length requirements – with annual observations one would need at least 20 years – as well as differing time periods, these methods come up with very divergent estimates about the size of Turkey’s unofficial economy. For instance using the “money demand” method Ögünç and Yılmaz (2000) report the “unofficial economy” to be between 0 and 46% of the official GDP for 1960-1998. At the other extreme Ilgin (2002) uses “tax collections” and 1985 to 2001 data. He finds the informal sector to be between 26 to 184% of Gross Domestic Product. Us’s (2004) estimate, based on electricity usage, ranges from 5 to 64% of GDP for the years 1997-2003.

Model based methods give explicit consideration to the multiplicity of causes underlying the unofficial sector. Thus they use multiple indicators to model these multiple causes, hence their acronym: MIMIC. Using this approach, Schneider and Savaşan (2005), estimate the size of the informal sector to be around 31-35% for the 1999-2005 period.

The method pioneered by Pissarides and Weber (1989)³, to be explained fully in our next section, falls in the indirect category. It essentially computes the marginal propensity to consume food using cross section data⁴. It extracts information about the magnitude of the informal sector by comparing the propensities of two or more groups. Arguably, it is preferable to the other, more macroeconomically oriented indirect methods. Firstly, the time series data used by the other methods cannot track year to year changes in the informal sector’s magnitude. For the mature economies of Western Europe, North America and Japan-Australia, the informal sector to GDP ratio can be taken to be stable. But for Latin America and the Middle East, such stability cannot be taken for granted. Moreover, for reasons discussed previously, the extent of informality in such countries is probably much greater than in the developed world. Secondly, Engel’s Law provides a theoretical framework to assess, independently albeit roughly, the accuracy of the estimates. This “Law” asserts that food expenditure to income ratio – or the average propensity to consume food and by implication the marginal one – falls as the income level rises. Therefore it implies the MPC for food of a “rich” country-period to be below that of a “poor” country-period. In Section 6, we will compare such estimates for a number of countries and show that our result passes this plausibility test, see Table 5.

There are a number of studies applying the P&W method to estimate the size of the “black economy” in various European countries. In addition to P&W (1989) which deals with the UK, there are Johansson (2005) on Finland, Schuetze (2002) on Canada and Engström & Holmlund (2006) on Sweden. It should be stressed all these papers focus on tax evasion by “self-employed” people like the “proverbial plumber.” Thus although they report high levels of income underreporting for this category, the overall size of the estimated “black economy” tends to be small. This follows from the low share of the “self employed” in total

³ Henceforth to be referred to as P&W.

⁴ One single year is enough.

employment for these advanced industrial democracies. For instance P&W (1989) find income underreporting by the self employed to be around 55%. Their GDP share being about 10%, this translates into a “black economy” of about 5.5% for the UK. Following a similar reasoning, Johansson reports the comparable figure for Finland to be about 2.5% of GDP. In countries like Turkey where tax evasion is not restricted to the “self-employed”, far greater magnitudes are involved. We now turn to this task.

4. Methodology

Three assumptions underlie P&W’s (1989) expenditure based estimation approach. They are:

- 1) All groups in the population report expenditure on *some items (especially food) accurately*.
- 2) The reporting of income by *some groups* in the population is *accurate*.
- 3) All groups in the population have the *same* expenditure pattern for *some items (like food)*.

Income underreporting is then modeled as follows: C_i represents food expenditure for household i on reported disposable income Y_i' . Z_i is a vector of household characteristics including the intercept. By assumption, every household reports C_i correctly. In this application Y_i' is correctly reported by people working in the “official/formal” economy⁵. Let Y_i stand for the “actual income” of household i . It follows that $Y_i = Y_i'$ for members of the formal sector. For those in the unofficial economy one has:

$$\text{Eq I: } Y_i = k_i Y_i', \text{ where } k_i \geq 1.$$

Here k_i is a random variable showing the extent of underreporting by household i . A bigger value indicates greater underreporting. For food expenditures, there is a “consumption function”,

$$\text{Eq II: } \ln C_i = \alpha Z_i + \beta \ln Y_i^p + \varepsilon_i. \text{ Here } \alpha \text{ is the parameter vector for household characteristics, } \beta \text{ measures the “marginal propensity to consume food” and } \varepsilon_i \text{ is white noise. } Y_i^p \text{ is the income measure that is valid for consumption decisions which is taken to be less volatile than disposable income } Y_i'. \text{ P\&W call this measure “permanent income”, without implying the conformity of Eq. II to the permanent income hypothesis.}$$

Actual and permanent incomes are related as follows:

$$\text{Eq III: } Y_i = p_i Y_i^p \text{ where } p_i \text{ is a random variable measuring the impact of aggregate shocks.}$$

Thus during “good”(“bad”) periods its mean will be above (“below”) unity. P&W assume that mean to be the *same* for both groups. However its variance is taken to be *larger* for members of the informal sector.

I and **III** together allow relating permanent income (Y_i^p) to reported disposable income (Y_i'):

$$\text{Eq. IV: } \ln Y_i^p = \ln Y_i' - \ln p_i + \ln k_i. \text{ Substituting IV into II yields:}$$

$$\text{Eq V: } \ln C_i = \alpha Z_i + \beta \ln Y_i' - \beta \ln p_i + \beta \ln k_i + \varepsilon_i.$$

We can now run the following regression:

$$\text{Eq VI: } \ln C_i = \alpha Z_i + \beta \ln Y_i' + \gamma \text{INF}_i + \eta_i. \text{ Here } \text{INF}_i \text{ is a binary variable equaling 1 when household } i \text{ is in the informal sector and 0 otherwise. The intercept is subsumed under } Z_i.$$

Keeping in mind our assumptions about p_i and k_i , namely the mean of p_i is the same for both groups, k_i is nonrandom and equals 1 for every member of the formal sector whereas $k_i \geq 1$ and is random for informal sector members, it can be seen that a *rough* estimate of income underreporting is given by γ/β or

⁵ In the original P&W (1989) paper, wage-salary receiving employees (as distinct from the self-employed) are assumed to report accurately.

Eq VII: $\ln k_i = \gamma/\beta$, see Figure 3⁶. For instance if $\gamma = 0.15$ and $\beta = 0.75$, their ratio equals 0.2. This means individuals working unofficially underreport their income by 20%. Alternatively, “average” underreporting can be expressed as $K = \exp(\gamma/\beta)$. Multiplying the reported disposable income by k , will give the true or actual disposable income. In this case K equals 1.221.

In Figure 3, $\ln C^*$ stands for the same accurately reported food expenditure level for two households, one formal the other informal. β gives the slope of both Engel curves whereas γ measures their vertical distance. Using a bit of trigonometry, we can see that their ratio (or $\ln k_i$) measures the horizontal distance between the two log income levels or the amount of underreporting. Of course this holds as a first approximation because it neglects the random nature of p_i and k_i which makes OLS estimation problematic. To solve this problem P&W assume these two to be log-normally distributed, and write them as deviations from their means. $\ln p_i = \mu_p + v_i$ and $\ln k_i = \mu_k + v_i$, where v_i and v_i have zero means and constant variances σ_v^2 and σ_v^2 within each group. Making the necessary replacements in Eq. IV and plugging it into II leads to

Eq. VIII: $\ln C_i = \alpha Z_i + \beta \ln Y_i - \beta(\mu_p - \mu_k) - \beta(v_i - v_i) + \varepsilon_i$.

A comparison of equations V and VIII shows that not only the intercept but also the error terms will be different for the two groups. Also informal group members will have larger error variances.

We need to estimate the “average” amount of underreporting or K . Using the log-normality assumption, P&W show that K can be expressed as:

Eq IX: $\ln K = \mu_k + \frac{1}{2} \sigma_{\text{vINF}}^2$ where σ_{vINF}^2 stands for the variance of k_i and INF refers to the informal sector members.

Using the normalized forms of p_i and k_i P&W argue that $\gamma = \beta \{ \mu_k + \frac{1}{2} (\sigma_{\text{vINF}}^2 - \sigma_{\text{vF}}^2) \}$ where σ_{vF}^2 stands for the variance of p_i for formal sector members. Plugging this expression in VIII results in:

Eq. X: $\ln K = \gamma / \beta + \frac{1}{2} (\sigma_{\text{vLNF}}^2 - \sigma_{\text{vINF}}^2 + \sigma_{\text{vF}}^2)$

Of course no data on p_i and k_i is available. Hence one cannot use this last equation to measure K . Instead P&W develop the following procedure to obtain upper and lower bounds on $\ln K$. One fits an income equation to get estimates for the income variance of errors for formal (σ_{YF}^2) and informal (σ_{YINF}^2) sector members separately. Using some simplifying assumptions they then derive the two bounds as:

Eq. XI: $\ln K = \gamma / \beta - \frac{1}{2} (\sigma_{\text{YINF}}^2 - \sigma_{\text{YF}}^2)$, lower bound.

Eq. XII: $\ln K = \gamma / \beta + \frac{1}{2} (\sigma_{\text{YINF}}^2 - \sigma_{\text{YF}}^2)$, upper bound.

5. Data

We used data from the Household Budget Survey (Hanehalkı Bütçe Anketi) for 2004 and 2005. Every year the Turkish Statistical Institute (TUIK, <http://www.tuik.gov.tr>) surveys 8600⁷ randomly selected households. The survey lasts one year and covers the whole country. Interviewees record each expenditure item for one month. They are asked to itemize their income (both cash and in kind) from every source (labor and non-labor). In addition the questionnaire contains numerous questions regarding household members’ employment and

⁶ Note that the two Engel Curves should emanate from the vertical axis, with the intercepts representing the impact of every factor except income.

⁷ Due to coding errors etc. the actual numbers are slightly less.

social security status, workplace characteristics. There are also questions covering marital and educational status as well as type of dwelling and the like. The results are used to estimate the distribution of disposable income among individuals and households. Another aim is to track the composition of household expenditures and its evolution through time. The results and other relevant information, such as sampling methodology – two stage clustered sampling – the questionnaires and all other pertinent material are available for a modest fee.

Following our previous discussion we can broadly define two types of informality: income underreporting in “registered/recorded” activities and “home production/moonlighting.” Typically, the income producing activities in this second category are not reported to⁸ or recorded by the tax authorities. In terms of our data we associate the first category with those cases where the household head works in an enterprise employing less than 50 workers⁹. Households headed by individuals who report no workplace are identified as “home producers.” Table 2 gives summary information on these points for both 2004 and 2005.

Since our sample is representative of the whole country, each observation comes with a weight expressing the number of households it represents. The sum of the weights gives the total number of households in the country, for instance 17.1 million in 2004. Multiplying a given year’s number of households figure with its average size, gives an estimate of the country’s population for that year, 72.5 million for 2005. It can be seen that more than 5 million households (roughly 30% of the total during both years), are headed by individuals who report “no workplace”. In assessing this rather high figure the following institutional quirk has to be kept in mind: up until 2002, the retirement age (in the formal sector) was 39 for women and 42 for men. Understandably, such “retirees” continue to work informally but hide it, in order not to lose their retirement benefits. Officially their number is about 2.2 million¹⁰. In addition, a sizable but falling number of such “no workplace” people dwell in rural areas and can be taken to be engaged in small farming. The last two rows give their breakup in terms of residential area¹¹. Comparing them gives an idea about their relative magnitudes. In 2004, roughly 27% of such (no workplace) households were urban dwellers; the next year their share goes up to 74%. The corresponding rural figures and ratios decline almost symmetrically. Even after allowing for sampling error, this reflects the considerable extent of rural to urban migration¹².

We now turn our attention to the income levels these groups report. According to our survey there are 17.1 million households in 2004. Of this figure, about 2.2 million are headed by individuals working in large firms employing more than 50 people. For 2005, the corresponding number is 2.4 million households. Based on the reasoning provided previously, we assume they report their income accurately both to the tax authorities and during the survey. Using the answers provided by the respondents, we calculate this category’s share of reported disposable income as: 16.9% in 2004 and 17.3% in 2005. As discussed in Sections 1 and 2, the rest of the population either does not turn in any income tax information or practices underreporting. In other words around 83% of reported disposable income represents the “informal economy.”

⁸ Businesses are required to file an income tax form. Employees are not. Supervision is lax.

⁹ In large (50+ workers) firms, according to TUIK (2005), only 6% of the workforce has no social security coverage. This percentage rises to 62-73 % in micro (10 – employees) enterprises where the bulk of employment takes place.

¹⁰ 15/2/2008 issue of Turkish daily Radikal page 15. Interview with Labor Minister Faruk Çelik.

¹¹ TUIK’s survey classifies an area with less than 20,000 inhabitants as rural.

¹² As pointed out by my discussant, the sampling error is large. For instance (after December 2007) TUIK reestimated the population in 2005 to be 70 instead of 72.5 million.

P&W method assumes that this category also underreports its income when responding to the expenditure survey, but reports its food spending accurately. We now present the results we get using their methodology.

6. Estimation Results

Table 3 gives some sample statistics. Formal sector members report higher incomes than their informal sector counterparts. Consistent with the P&W hypothesis informal sector people

have considerably higher income variance. Also a Kolmogorov-Smirnov test for log-normality of income strongly supports the null hypothesis. Apart from these, formal sector members seem to spend slightly more than their informal sector counterparts. They are also younger and better educated. On the other hand homeownership is more prevalent among informal sector members who tend to be older. These are all consistent with a priori expectations and the previously mentioned institutional quirk involving the retirement age.

Table 4 displays the heteroskedasticity corrected β and γ estimates obtained by instrumental variables (IV) and ordinary least squares methods. The “marginal propensity to consume” food estimate is around 0.72 for both years. The coefficient for the informal sector (INF) dummy shows some small variation across the two years but averages around 0.15. Taking the antilog of their ratio gives $\{\exp(.15/.72)\}$ a rough estimate of K to be 1.23. The reported disposable income for the informal sector should be multiplied by this figure to get their actual disposable income. Before computing the lower and upper bounds of this multiplier using equations 11 and 12, we do a plausibility check of our MPC for food estimate. We compare our 0.72 to those reported in the literature and check them against Engel’s Law. This well-known “law” predicts that as income levels rise, the average (and therefore marginal) propensity to consume food will fall. Table 5 displays the relevant magnitudes – in addition to Turkey– for Canada, Finland, Sweden and UK. They are from Schuetze (2002), Johansson (2005), Engström & Holmlund (2006) and Pissarides & Weber (1989) respectively. The real per capita GDP figures are from Penn World Tables. They are “constant” magnitudes which are comparable across time and space. The numbers in parenthesis give the relevant years¹³.

It can be seen that consistent with Engel’s Law, Turkey has the lowest income level and the highest marginal propensity to consume food. This table provides some independent confirmation that our finding is sensible.

Finally equations 11 and 12, together with the income variances, lead to the following lower and upper bounds:

Lower Bound, 2004: $(0.14/0.73) - \frac{1}{2} (0.762 - 0.677) = 0.150$; $K_{low} = \exp(0.150) = 1.16$

Upper Bound, 2004: $(0.14/0.73) + \frac{1}{2} (0.762 - 0.677) = 0.235$; $K_{up} = \exp(0.235) = 1.26$

The corresponding calculations for 2005 are:

Lower Bound, 2005: $(0.162/0.725) - \frac{1}{2} (0.792 - 0.666) = 0.160$; $K_{low} = \exp(0.160) = 1.17$

Upper Bound, 2005: $(0.162/0.725) + \frac{1}{2} (0.792 - 0.666) = 0.287$; $K_{up} = \exp(0.287) = 1.33$.

7. Summary

The P&W approach is based on two premises. Firstly, all segments of the population correctly report their food expenditures when surveyed. Secondly, informal sector members hide some/all their income, but members of the formal sector report their income accurately. Using these assumptions and the model displayed in equations 1 through 12, we estimate food expenditure equations for members of both sectors. Using the logic displayed in Figure

¹³ Note that Penn World Table figures –which are comparable across time and space– stop at 2000.

3 and equations 11 and 12, we conclude that for 2005, reported informal sector income, should be multiplied by about 1.25, to obtain actual income. From the Household Budget Survey figures we know the share of total (reported) disposable income for this group to be 83%. It follows that according to our estimate, true disposable income is $(0.25 \cdot 0.83)$ about 21% larger than the reported one. The other most reliable estimate is 31 to 35% of GDP by Schneider and Savaşan (2005). Although smaller, ours is in the same order of magnitude. The difference might be due to the “disposable” versus “gross domestic” distinction, as well as the time periods involved.

Appendix:

We used the following variables (* denotes element of household characteristics vector Z ; ** indicates instrument to estimate LNINC).

LNFOOD = natural log of food expenditures

INF = informal sector dummy equaling 1 for sector members

INDEP = dummy equaling 1 for independent business owners (*)

EDUC = years of schooling (*)

HHSIZE = number of household members (*)

ROOMS = number of rooms in the house (*)

CAR = dummy equaling 1 for car owners (*)

LNINC = natural log of reported disposable income

AGE = head of household's age (**)

AGESQ = age squared (**)

GENDER = household head's gender (**)

RENT = rent paid – imputed for homeowners. (**)

HOME = home ownership dummy 1 means owner (**)

URBAN = area dummy; 1 for urban dwellers. (**)

Estimation was performed using Stata 9.1's ivreg2 procedure. We used the heteroskedasticity corrected GMM method. Table 6 displays the full results.

References:

- Adaman, F; A. Çarkoğlu and B. Şenatalar (2001). “Hanehalkı Gözünden Türkiye’de Yolsuzluğun Nedenleri ve Önlenmesine İlişkin Öneriler.” TESEV, İstanbul. (In Turkish)
- Al-Kawaz (1999). *Informal Sector and Its Role in Arab Countries*. American University in Cairo Press, Cairo.
- Besim, M and G. Jenkins (2005). “Tax Compliance: When Do Employees Behave Like the Self-employed?” *Applied Economics* 37, 1201-1208.
- Derviş, K., D. Gros, M. Emerson and S. Ülgen (2004). *Çağdaş Türkiye’nin Avrupa Dönüşümü*. Doğan Kitap, İstanbul. (In Turkish)
- Djankov, S., I. Lieberman, J. Mukherjee and T. Nenova (2002). “Going Informal: Benefits and Costs.” *Decent Work and the Informal Economy*. Report VI, World Bank-ILO Conference, 90th Session, Geneva.
- Engström, P. and B. Holmlund (2006). “Tax Evasion and Self Employment in a High Tax Country: Evidence from Sweden.” *CESIFO Working Paper No. 1736*, www.CESifo-group.de
- Henry, C.M and R. Springborg (2001). *Globalization and the Politics of Development in the Middle East*. Cambridge University Press
- Ilgın, Y. (2002). “Kayıt Dışı Ekonomiyi Tahmin Yöntemleri ve Türkiyede Durum.” *DPT Planlama Dergisi*, Özel Sayı. (In Turkish)
- Johansson, E. (2005). “An Estimate of Self-Employment Income Underreporting in Finland.” *Nordic Journal of Political Economy* 31, 99-110.
- OECD (2004). *Economic Survey of Turkey*, OECD, Paris. http://www.oecd.org/document/44/0,2340,en_33873108_33873854_33804140_1_1_1_1_00.html Penn World Tables: <http://pwt.econ.upenn.edu/>
- Öğünç, F. and G. Yılmaz (2000). “Estimating Underground Economy in Turkey.” Central Bank, Republic of Turkey, Discussion Paper.
- Pissarides, C and G Weber (1989). “An Expenditure-Based Estimate of Britain’s Black Economy.” *Journal of Public Economics* 39, 17-32.
- Sapir, A (2005). “Globalisation and Reform of European Social Models.” Bruegel Policy Brief No 1, http://www.bruegel.org/doc_pdf_120, Brussels.
- Schneider F. and D. Enste (2000). “Shadow Economies: Size, Causes and Consequences.” *Journal of Economic Literature* 38: 77-114.
- Schneider, F. and F. Savaşan (2005). “The Size of Shadow Economies of Turkey (and of her Neighbouring Countries) Including an Informal Hiring and Sectoral Analysis of the Turkish Shadow Economy.” <http://www.economics.uni-linz.ac.at/Schneider>
- Schuetze, H (2002). “Profiles of Tax Noncompliance Among the Self-Employed in Canada: 1969-1992.” *Canadian Public Policy* 28, 219-238.

TUIK (2005). “Hane Halkı İşgücü Anketi.” (In Turkish)

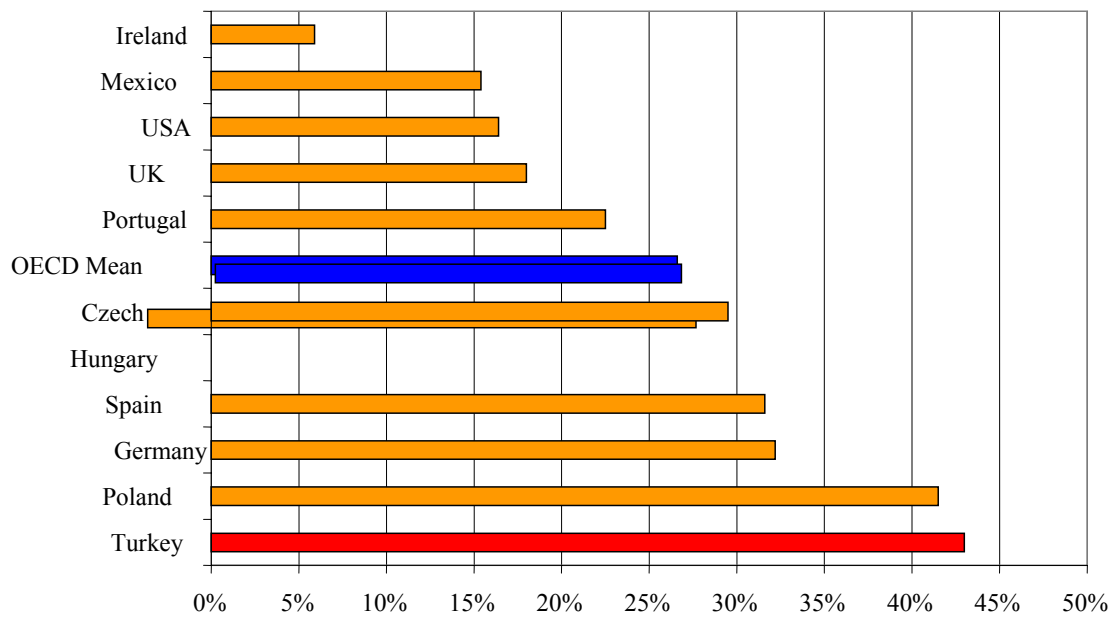
Ulgen, S and U. Ozturk (2006). “AB Surecinde Türkiye’de Kayıtdışı Ekonomi.” (In Turkish)<http://www.tusiad.org/turkish/rapor/buyume/kayitdisi.pdf>

Us, V. (2004). “Kayıtdışı Ekonomi Tahmini Yöntem Önerisi: Türkiye Örneği.” *Türkiye Ekonomi Kurumu*, 2004/17 (In Turkish)

World Bank (2006). Turkey Labor Market Study, Report No 33254-TR.

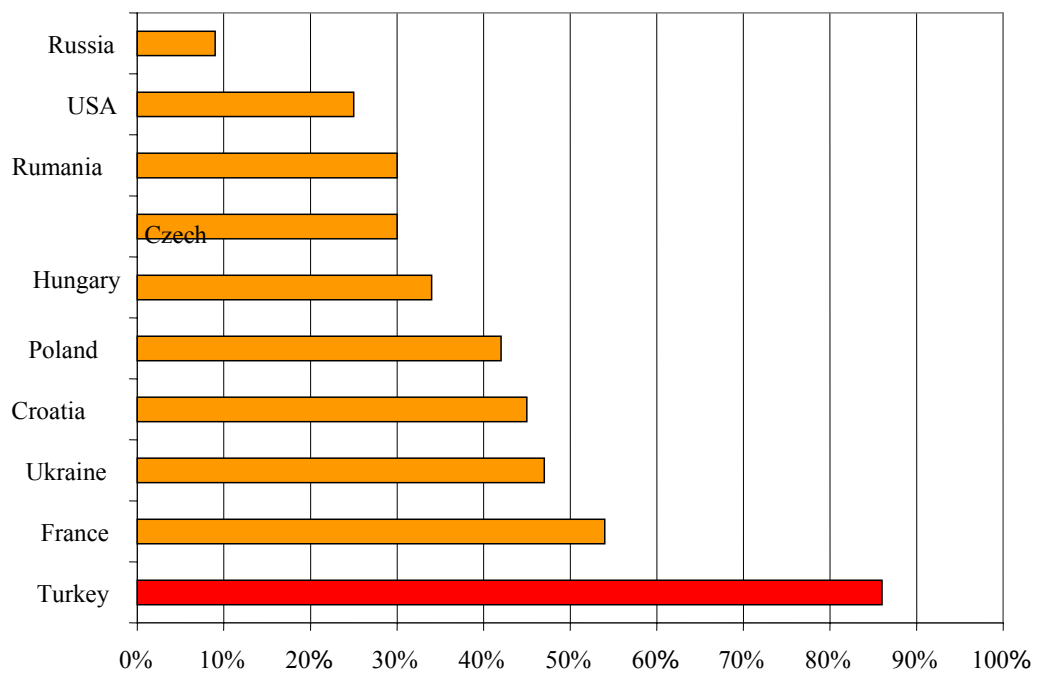
World Bank Doing Business Database: www.doingbusiness.org.

Figure 1: Burden of Employment Taxes in OECD Countries (2004)



Source : OECD Taxing wages 2003-2004, February 2005

Figure 2: Minimum Wage to Per Capita GDP Ratio (2006)



Source : IMF World Economic Outlook Database April 2006, CIA World Factbook: yearly average dollar exchange rates.

Figure 3: Engel Curves for Members of FORMAL and INFORMAL Sectors

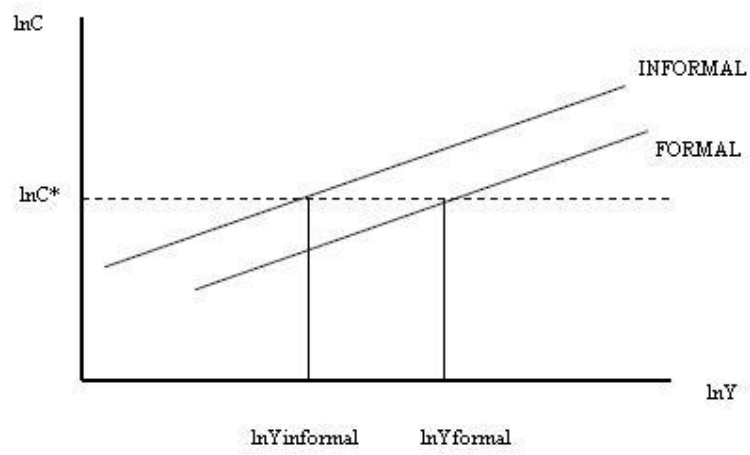


Table 1: Data on Employment Policy: Hiring & Firing (2006)

Country	Hiring Difficulty Index	Firing Difficulty Index	Hiring Cost (% Monthly Compensation)	Firing Cost (#of Weekly Wages)
OECD Average	33	32	26	39
Czech	33	30	35	22
France	67	40	47	32
Germany	33	40	19	69
Greece	44	50	31	69
Hungary	11	10	35	35
Italy	61	40	42	2
Japan	28	0	13	9
Mexico	33	40	24	74
Poland	0	40	21	13
Portugal	33	60	24	99
Spain	78	50	30	56
UK	11	10	11	22
USA	0	0	9	0
Turkey	56	30	22	95

Source: World Bank Doing Business

Table 2: Survey Summary (Weighted Numbers in Parenthesis)

	2004	2005
Total # of households in the sample	8544 (17,100,000)	8559 (17,500,000)
Average household size	4.141	4.147
Household head's workplace employs less than 10 People	3,583 (7,399,562)	3,626 (7,571,758)
10 to 24	792 (1,604,183)	751 (1,584,605)
25 to 49	406 (842,133)	375 (715,565)
Household head's workplace employs 50+	1,148 (2,203,892)	1,220 (2,374,951)
No workplace information	2,615 (5,046,768)	2,587 (5,302,142)
No workplace information & Rural dwellers	2,049 (3,667,883)	546 (1,360,220)
No workplace information & Urban dwellers	566 (1,378,885)	2,041 (3,941,922)

Table 3: Sample Statistics

	2004		2005	
	Formal	Informal	Formal	Informal
Number	1148	7396	1220	7339
Age	38.83 (7.66)	48.25 (13.72)	39.13 (7.93)	48.44 (13.87)
Education in years	7.31 (2.96)	5.80 (2.86)	5.88 (2.63)	4.06 (2.39)
Household Size	4.05 (1.27)	4.16 (2.12)	4.11 (1.53)	4.15 (2.12)
Homeowner	0.61 (0.48)	0.80 (0.40)	0.63 (0.48)	0.79 (0.41)
lnY(income)	16.89 (0.87)	16.54 (0.95)	17.06 (0.81)	16.7 (0.98)
Ln C (food consumption)	12.77 (0.78)	12.73 (0.84)	12.91 (0.77)	12.88 (0.84)
Kolmogorov-Smirnov (K-S) Test of normality for lnY	0.0182 (0.833)	0.0092 (0.557)	0.016 (0.909)	0.0075 (0.794)

Note: The numbers in parenthesis are the standard errors, except for the K-S statistics where they represent p-values. A larger p-value favors the null of normality.

Table 4: β and γ Estimates (t values in Parenthesis)

	B	γ	γ/β	K	σ_{YINF}^2	σ_{YF}^2
IV 2004 (N=8541)	0.730 (43.87)	0.140 (8.45)	0.192	1.16	0.762	0.677
IV 2005 (N=8550)	0.725 (42.14)	0.162 (9.70)	0.223	1.25	0.792	0.666
OLS 2004 (N=8541)	0.717 (95.13)	0.147 (8.99)	0.205	1.23		
OLS 2005 (N=8550)	0.710 (90.05)	0.152 (9.29)	0.214	1.24		

Table 5: Consistency with Engel's Law

	Turkey	Canada	UK	Finland	Sweden
MPC food ¹⁴ (period)	0.720 (2004)	0.389 ¹⁵ (1969-92)	0.253 (1982)	0.249 (1994-96)	0.246 (1999-2004)
Income (period)	6,838 (2000)	14,106 (1969)	14,412 (1982)	18,789 (1995)	23,661 (2000)

¹⁴ When both IV and OLS estimates are reported, their average is used. The UK figure is the blue collar, white collar average.

¹⁵ The author constrains β to be equal across the whole period. The pure 1969 estimate would have been higher.

Table 6: IV and OLS Estimates (t Values in Parenthesis)

Dep. Var LNFOOD	2004 (N= 8541)		2005 (N=8550)	
	IV	OLS	IV	OLS
Constant	0.582 (2.27)	0.808 (6.68)	0.736 (2.73)	0.997 (7.66)
LNINC	0.730 (43.87)	0.717 (95.13)	0.725 (42.14)	0.710 (90.05)
INF	0.140 (8.45)	0.147 (8.99)	0.162 (9.70)	0.152 (9.29)
INDEP	-0.179 (-8.40)	-0.154 (-7.56)	-0.127 (-6.12)	-0.132 (-6.62)
EDUC	-0.042 (-14.60)	-0.041 (-16.77)	-0.047 (-15.85)	-0.045 (-18.46)
HHSIZE	0.064 (21.06)	0.064 (21.34)	0.065 (21.22)	.066 (21.22)
ROOMS	-0.021 (-2.38)	-0.025 (-3.00)	-.027 (-3.19)	-0.028 (-3.40)
CAR	-0.056 (-3.86)	-0.065 (-4.55)	-0.105 (-6.88)	-0.106 (-7.58)
Constant	14.58 (128.58)		15.26 (131.45)	
AGE	0.21 (5.07)		0.036 (8.96)	
AGESQ	-0.00018 (-4.57)		-0.00031 (-7.77)	
GENDER	-0.176 (-6.31)		-0.186 (-6.03)	
RENT	3.18e-09 (30.06)		0.0027 (17.93)	
HOME	0.136 (6.24)		0.080 (3.56)	
URBAN	0.330 (16.00)		-0.293 (-11.78)	
R ² LNFOOD	0.591	0.592	0.584	0.585
R ² LNINC	0.370		0.355	