Reducing the Shadow Economy through Electronic Payments TECHNICAL APPENDICES

This study was commissioned by Mastercard and was conducted independently by EY.

EY
Building a better working world

Table of contents

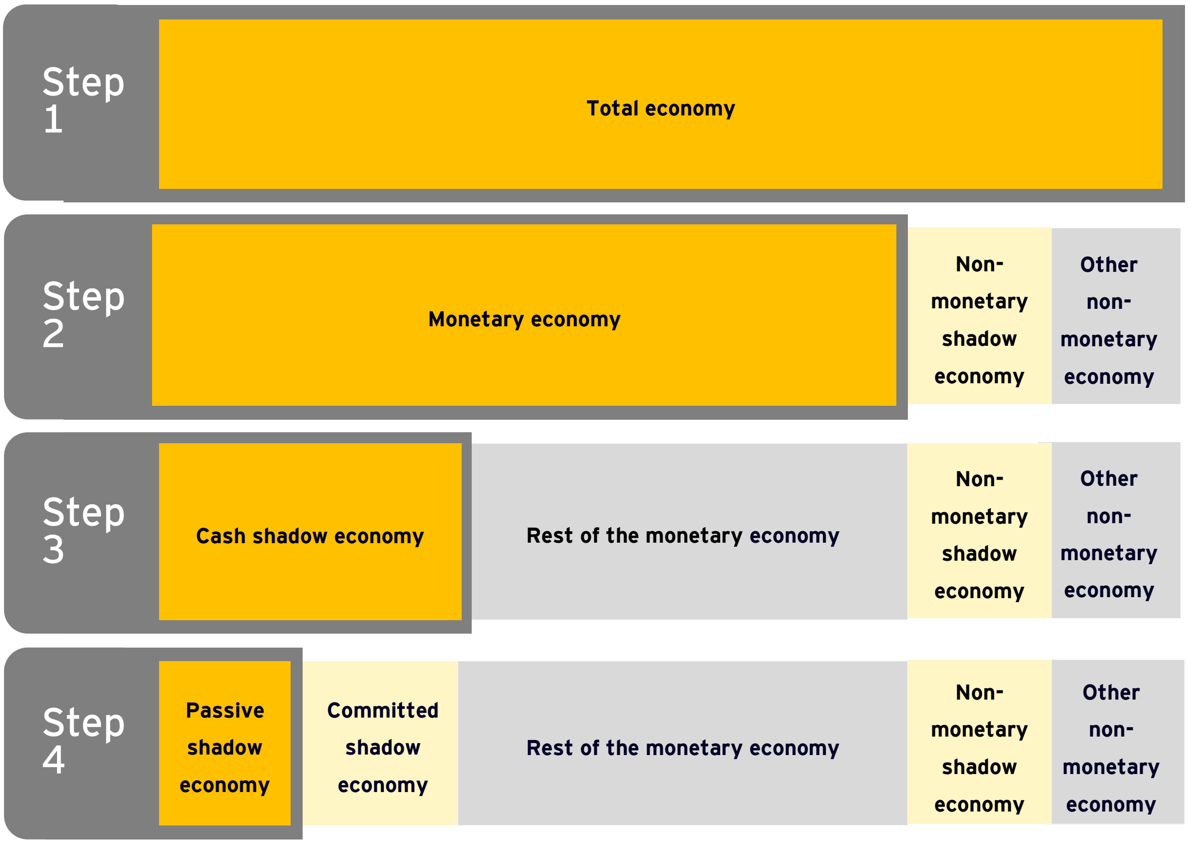
Introduction

The document "Reducing the Shadow Economy through Economic Payments: TECHNICAL APPENDICES" is part of the broader material that also consists of individual reports "Reducing the shadow economy through electronic payments" for a broad group of countries (hereinafter referred to as the "Reports"). The Reports analyse the shadow economy in a number of countries and investigate the potential of different solutions to reduce the size of the non-observed economy. The current document comprises the technical appendices that provide more details on our approach to the measurement of the shadow economy and its decomposition, as well as to estimating the effects of various regulatory measures. More information and documents are available on: http://www.ey.com/pl/electronic-payments.

Appendix 1. Estimation of the shadow economy: methodological details

In this appendix, we discuss in detail the methodology applied for estimating the size of the shadow economy and present additional results which are not included in the Reports. The structure of the appendix follows our general approach towards estimating the structure and size of the shadow economy, in which we start with defining the total economy (step 1), next we carry out a decomposition that distinguishes the monetary economy and the non-monetary economy (including the non-monetary shadow economy) (step 2), the monetary economy is then broken down into the cash shadow economy and the rest of the monetary economy (step 3), and, finally, we distinguish the passive and committed components of the cash shadow economy, which is crucial for our analysis (step 4; see **Chart A1.1**).

Chart A1.1. Decomposition of the total economy into shadow and official components



The proportions of the areas above do not reflect the proportions of different components of the total economy.
Source: EY elaboration.

Step 1. Determining the relation between the total economy and the official GDP

Our methodology comprises the estimation of different shadow economy components only as percentage of the size of the total economy (total GDP[1]). To express these estimates in local currency units or as a percentage of official GDP figures, we need to know first the relation between the size of the total economy and the official GDP figures published by the statistical offices ($\frac{Y_{i,t}^{TOTAL}}{Y_{i,t}^{OFFICIAL}}$ for country i in period t).

[1] GDP (gross domestic product) is a monetary measure of the market value of all final goods and services produced in a given period of time on a given territory. It is closely related to the notion of gross value added (or simply value added) that can be used to measure the production in a given sector (e.g. agriculture, construction, etc.).

We start with analysing the scope of the official GDP figures in terms of their coverage of shadow (non-observed) economy activities. There are two options. First, we may know the size of the non-observed economy included in official GDP figures (it could be equal to zero), e.g. from information included in statistical offices publications or from our correspondence with these institutions. In such a case, if our estimate of the share of the shadow economy in the total economy is different than the share of the statistical office's estimate of the shadow economy in official GDP, it means that the total economy differs from official GDP figures. To account for these differences we run necessary recalculations to present the results of our estimations in terms of percentage of official GDP figures (e.g. see substep 3.3 for a discussion of the procedure for the cash shadow economy).

Second, there may be no statistical offices' estimates of the shadow economy available[2]. Despite this, the official GDP figures usually should include unregistered activities, in compliance with (i) the national accounts guidelines applied by a given statistical office (e.g. System of National Accounts 2008 worked out by the United Nations) or (ii) the methodology used by the statistical offices to estimate the value added in different sectors.[3] In such a case, we assume that our estimates of the shadow economy are equal to the unregistered value added included in official GDP. Consequently, there is no need to conduct any recalculations.

The recalculation from the fraction of total GDP to the fraction of officially registered GDP applies not only to the cash shadow economy, but also to other shadow economy measures discussed in the Reports, i.e. the non-monetary shadow economy, as well as the passive and committed components of the cash shadow economy.

Step 2. Measuring the monetary economy and estimating the non-monetary shadow economy

The total economy comprises both monetary, i.e. payment-based, activities and non-monetary activities, unrelated to any monetary payments. In this step of the analysis, we distinguish the monetary and non-monetary parts of the economy and decompose the latter into shadow and non-shadow segments (see **Chart A1.1**).

It is worthwhile to mention that the leading existing analytical method of shadow economy measurement, i.e. the currency demand approach (CDA), is directly applicable to the monetary economy only. Consequently, unless additional recalculations are conducted, the obtained shadow economy estimates should be expressed in terms of percentage of the total monetary economy. However, in literature these estimates are (usually) expressed in percentage of the total GDP or, even worse, official GDP. Such an approach implicitly assumes that there is no value added generated through non-monetary activities in a given economy, which may lead to a significant overestimation of the presented shadow economy estimates.

By contrast, in our approach we account for both the fact that official GDP figures often include at least some shadow economy estimates (see step 1) and that the size of the total economy exceeds the size of the monetary economy. To our knowledge, such an approach has not been applied in the previous analyses of the cash shadow economy, so we consider it as our contribution to the existing literature.

To estimate the share of the monetary economy in the total economy (total GDP), we subtract two significant components of the non-monetary economy from the total economy: 1) the value of

[2] A statistical office might even be unable to identify the share of the shadow economy included in the GDP figures published by the office itself, since GDP figures are computed within a complex procedure based on data from various sources and their reconciliation, rather than as an aggregate of shadow- and non-shadow components.

[3] For example, in some countries the non-observed economy is likely to be included in the case of agriculture (and household production of goods for own consumption which is mostly related to agriculture), since its value added is calculated on the basis of the harvested area and the assumed average productivity (value added) per area unit (e.g. for a given type of crop).

imputed rents[4] that are included in national accounts and could be found in statistical offices datasets and 2) the value of household production of goods for own final use.[5] The latter component is further referred to as "the non-monetary shadow economy" (see also Chapter 1.1 in the Reports and **Chart A1.1**).[6]

Wherever possible, we derive the estimates of the value of household production of goods for own final use directly from the national accounts datasets. However, this is a rare case, even though most countries include this kind of production in the estimates of value added for different sectors and GDP figures if it is material (especially in the case of agriculture, see also footnote 4). Otherwise, the size of the non-monetary shadow economy may be estimated on the basis of the existing, though limited, research in this area. For this purpose we use the results of Blades (1975)[7] who conducted an extensive survey of the size of household production of goods for own final use in various countries, especially developing ones.

We estimate the size of the non-monetary shadow economy for a given country and year by using (i) data on the share of agriculture in GDP and (ii) the non-linear equation estimated by Blades that shows the relation between the share of agriculture in GDP and the share of the non-monetary shadow economy in GDP (see the grey line on **Chart A1.2**). This relation results from the fact that a significant part of the non-monetary shadow economy covers the agricultural production. Since the Blades equation was estimated as non-linear for countries with a relatively high share of agriculture in GDP, it may suffer from a high out-of-sample prediction error for the countries in which the role of agriculture and/or non-monetary shadow economy is low. In particular, this equation indicates that the non-monetary shadow economy is significant in the proximity of zero agriculture share as the equation was fit on the sample of countries with relatively high agriculture shares. Instead, we may assume that in a hypothetical highly developed country in which the share of agriculture in GDP is equal to zero, the share of the non-monetary shadow economy in GDP should be also close to zero. To address this problem, for countries with a low share of the non-monetary shadow economy in GDP (below 10%), we replace the equation of Blades within this value range with a linear dependency, starting in the point (0;0) and sloped positively so as to keep the function continuous (see the yellow line on **Chart A1.2**).

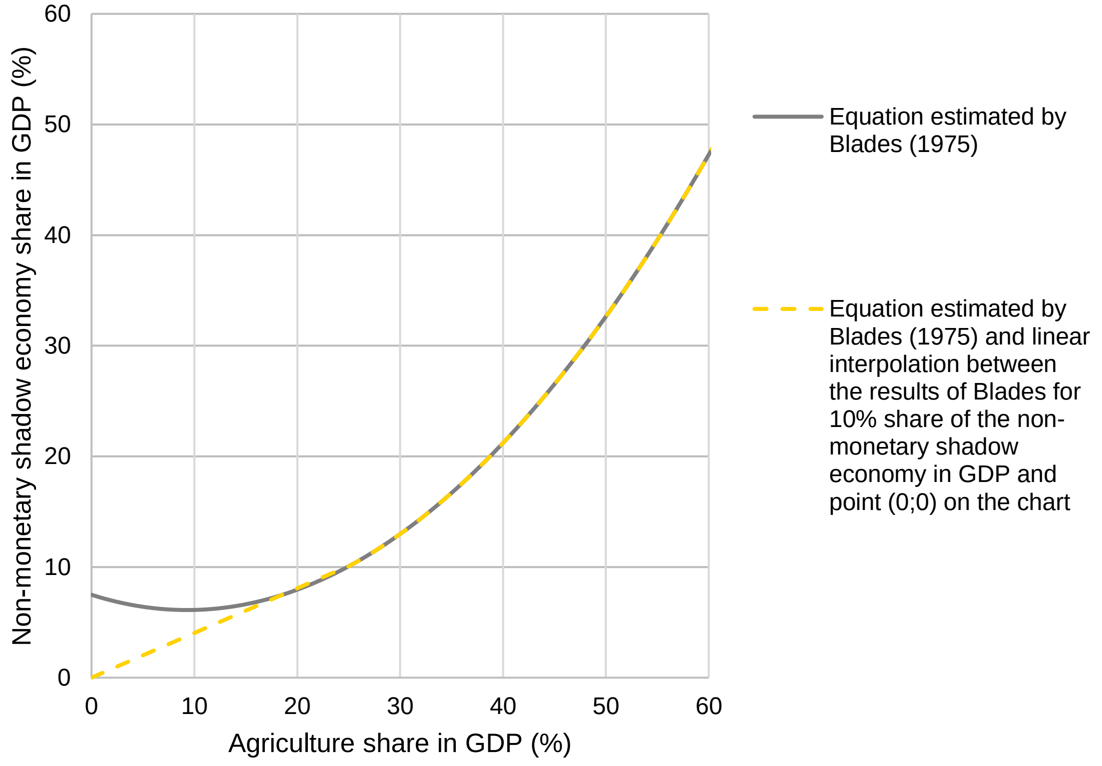
[4] Imputed rents are related to housing services that homeowners implicitly provide for themselves. They are estimated to be equal to the rents that homeowners would have paid to live in dwellings of the same type, in the same district and with the same service facilities. They are included in GDP figures. If they were not, the GDP would be affected by changes in the share of people living in their own dwellings. It is assumed that, for example, a situation in which two homeowners living in their own dwellings start letting their dwellings to each other and paying regular rents should not affect the level of GDP. Indeed, such a change does not impact the level of GDP, since these "new" rents have already been included in GDP as imputed rents.

[5] We do not account for household services for own final use (e.g. cooking for a spouse) since they are generally excluded from national accounts statistics, including GDP, with an exception of the imputed rents mentioned above.

[6] Household production of goods for own consumption (not sold on the market) could be treated as a part of the non-observed economy, which is non-monetary (see, e.g., OECD (2002), "Measuring the Non-Observed Economy. A Handbook"). This is also the case in our study. According to national accounts guidelines, the estimated value of such production should be included in GDP, especially if it accounts for a significant part of the total supply of a given good in a country.

[7] Blades D. W. (1975), "Non-Monetary (Subsistence) Activities in the National Accounts of Developing Countries", Organisation for Economic Co-operation and Development, Paris.

Chart A1.2. Estimation of the non-monetary shadow economy size based on agriculture share in GDP



Source: EY elaboration based on Blades D. W. (1975), "Non-Monetary (Subsistence) Activities in the National Accounts of Developing Countries", Organisation for Economic Co-operation and Development, Paris.

Step 3. Measuring the cash shadow economy (currency demand analysis – CDA)

CDA: the general principle

In step 3 we estimate the share of the shadow economy in the official economy ($\frac{Y_{i,t}^{TOTAL,SHADOW}}{Y_{i,t}^{OFFICIAL}}$). To achieve this we first focus on measuring the share of the monetary (or "cash") shadow economy in total monetary economy ($\frac{Y_{i,t}^{CASH,SHADOW}}{Y_{i,t}^{MONETARY,TOTAL}}$; see **Chart A1.1**).

The currency in circulation (cash) data, which is collected by statistical offices or central banks, conveys useful information about both registered and non-registered monetary economic activities. Accordingly, the demand for cash is often decomposed econometrically within the currency demand analysis into the two components:

- **structural demand for cash** – reflecting the need for certain amount of cash to be used in normal, registered economic activities (for instance, people hold cash for precautionary or liquidity purposes);
- **"excessive" demand for cash** – related to shadow economy activities.

In CDA, it is assumed that the share of an "excessive" cash component in total transactional money (in the form of cash and demand deposits) is equal to the share of the cash shadow economy in the monetary economy.

The key advantage of this approach is that it allows us to estimate the level of the cash shadow economy and its development over time with a very limited need for additional, external information. In addition, as our study focuses on the reduction of the shadow economy by replacing cash with electronic payments, the direct relation of CDA to the use of cash in the economy is yet another advantage of this approach.

However, the traditional CDA framework has also some limitations. Firstly, it does not allow us to distinguish the committed and passive components of the cash shadow economy (see Chapter 1 in each of the Reports), which play a critical role in measuring the effects of solutions promoting electronic payments. Secondly, it is not straightforward to take account of the factors that affect both components of cash demand simultaneously, i.e. the structural and the "excessive" component, such as the popularity of electronic payments. Thirdly, a crucial step of the CDA-based decomposition of the cash demand is to compute its theoretical value in a hypothetical situation where the shadow economy does not exist. To this aim, many authors unrealistically set shadow-economy-related variables to zero, for example the tax rate, which may lead to an upward bias in the results.[8]

Inspired by the existing CDA research[9], but also accounting for its weaknesses, we propose a modified approach, described in more detail in an academic article prepared by the authors of the Reports[10] and already recognized by other shadow economy researchers.[11] In our approach we distinguish four substeps, which we describe below.

Substep 3.1. Estimation of CDA model

The first substep consists in an econometric estimation of the currency demand equation:

$$\boldsymbol{CASH\_M1}_{i,t} = \alpha_i + \boldsymbol{\beta}^{(1)}_{i,t}\boldsymbol{x}_{\mathbf{1},i,t} + \boldsymbol{\beta}^{(2)}_{i,t}\boldsymbol{x}_{\mathbf{2},i,t} + \boldsymbol{\beta}^{(3)}_{i,t}\boldsymbol{x}_{\mathbf{3},i,t} + \varepsilon_{it}, \text{(A1.1)}$$

where i represents the analysed country and t stands for the analysed time period (we use quarterly data). In this equation, the explained (dependent) variable is the share of currency in circulation ("cash") in the M1 monetary aggregate ("total transactional money"). To explain its variation, we use three groups of explanatory variables (see **Table A1.1** for details and sources):

- **"Typical" cash shadow economy determinants ($\mathbf{x_1}$).** The impact of these variables on the dependent variable is mostly related to the incentives and disincentives of the economic agents to operate in the shadow economy (see Chapter 1 of the Reports).
- **Payment card system variables ($\mathbf{x_2}$).** Higher levels of these variables (reflecting better development of the payment card system) may be associated with two effects: (1) a decrease in the size of the shadow economy (by reducing the opportunities for leaving transactions unreported) and (2) replacement of the registered cash transactions with electronic payments (no impact on the size of the cash shadow economy).
- **Other control variables ($\mathbf{x_3}$).** These variables, after controlling for the influence of $\mathbf{x_1}$ and $\mathbf{x_2}$ determinants, should not (directly) impact on the size of the shadow economy, but may still have some influence on the dependent variable. These variables are related to the level of the economic development, technical progress and institutional factors (some of which might be unobservable and must therefore take the form of dummy variables).

$\boldsymbol{\beta}^{(1)}_{i,t}$, $\boldsymbol{\beta}^{(2)}_{i,t}$ and $\boldsymbol{\beta}^{(3)}_{i,t}$ represent vectors of the regression coefficients. We assume that these coefficients depend on real GDP *per capita* (Purchasing Power Parity adjusted, US dollar in 2011), so they differ across countries and periods (which is reflected in the notation by i and t superscripts). For $k=1,2,3$

---

[8] The zero tax approach (or zero incentives to move to shadow economy approach) was used by Tanzi V. "The Underground Economy in the United States: Annual Estimates, 1930-80", Staff Papers – International Monetary Fund, Vol. 30, No. 2, 1983, pp. 283-305. For more recent examples, see Embaye A., "Underground economy estimates for non-OECD countries using currency demand method, 1984-2005," MPRA Paper 20308, University Library of Munich, Germany, 2007 and Ardizzi G., Petraglia C., Piacenza M., Turati G., "Measuring the underground economy with the Currency Demand Approach: A reinterpretation of the methodology, with an application to Italy", *Review of Income and Wealth*, vol. 60(4), 2014, pages 747-772. This approach is also suggested by Ahumada H., Alvaredo F., Canavese A. J. (2006), "The Demand for Currency Approach and the Size of the Shadow Economy: A Critical Assessment", *Berkeley Program in Law & Economics, Working Paper Series*.

[9] See Tanzi (1983), Ardizzi et al. (2014), *op. cit*.

[10] Dybka P., Kowalczuk M., Olesiński B., Rozkrut M., Torój A. (2018), "Currency demand and MIMIC models: towards a structured hybrid method of measuring the shadow economy", *International Tax and Public Finance*, pages 1-37, https://link.springer.com/article/10.1007/s10797-018-9504-5 [online; accessed 13.07.2018].

[11] See e.g. Medina L., Schneider F. (2018), "Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?", *IMF Working Paper*, no. WP/18/17.

we have:

$$\boldsymbol{\beta}_{i,t}^{(\boldsymbol{k})} = \boldsymbol{\gamma}_{\mathbf{0}}^{(\boldsymbol{k})} + \boldsymbol{\gamma}_1^{(\boldsymbol{k})} \boldsymbol{GDP\_REAL\_PER\_CAPITA_{it}}, \text{ (A1.2)}$$

in which $\boldsymbol{\gamma}_{\mathbf{0}}^{(\boldsymbol{k})}$ and $\boldsymbol{\gamma}_1^{(\boldsymbol{k})}$ are vectors of parameters describing coefficient heterogeneity. Essentially, some coefficients out of $\boldsymbol{\beta}_{i,t}^{(\mathbf{1})}$, $\boldsymbol{\beta}_{i,t}^{(\mathbf{2})}$, $\boldsymbol{\beta}_{i,t}^{(\mathbf{3})}$ can be constant. In such cases, the respective element of the vector $\boldsymbol{\gamma}_1^{(\boldsymbol{k})}$ equals 0, i.e., the strength of the coefficient out of $\boldsymbol{\beta}_{i,t}^{(\mathbf{1})}$, $\boldsymbol{\beta}_{i,t}^{(\mathbf{2})}$, $\boldsymbol{\beta}_{i,t}^{(\mathbf{3})}$ does not vary with GDP *per capita*. To estimate $\boldsymbol{\gamma}_1^{(\boldsymbol{k})}$, we use the interaction terms, i.e. products of the respective variables with $\boldsymbol{GDP\_REAL\_PER\_CAPITA_{it}}$. Finally, ε_{it} is the error term.

Additionally, we include the individual effects, α_i, which represent time-invariant, unobservable country characteristics that affect the demand for cash in each country.

The construction of the coefficients α_i, $\boldsymbol{\beta}_{i,t}^{(\mathbf{1})}$, $\boldsymbol{\beta}_{i,t}^{(\mathbf{2})}$ and $\boldsymbol{\beta}_{i,t}^{(\mathbf{3})}$ reflects country heterogeneity. This is crucial in our econometric model because to estimate the parameters we use data for 64 countries[12]: Albania, Algeria, Angola, Argentina, Armenia, Australia, Azerbaijan, Bahrain, Bangladesh, Bolivia, Bosnia and Herzegovina, Brazil, Bulgaria, Canada, Chile, China (Mainland), Colombia, Croatia, Czech Republic, Denmark, Dominican Republic, Egypt, Honduras, Hungary, India, Indonesia, Israel, Jamaica, Japan, Jordan, Kazakhstan, Kuwait, Lebanon, Malaysia, Mexico, Moldova, Mongolia, Myanmar, Nepal, New Zealand, Nigeria, Norway, Pakistan, Peru, Philippines, Poland, Romania, Russia, Rwanda, Saudi Arabia, Serbia, Singapore, Sri Lanka, Sweden, Switzerland, Tanzania, Thailand, Tunisia, Turkey, Ukraine, United Arab Emirates, United Kingdom, Uruguay, Vietnam. The sample consists of the available quarterly observations for these countries over the years of 2000-2016 and effectively covers different time periods for different countries (unbalanced panel). Some variables that have not been available at a quarterly frequency have been linearly interpolated into quarterly data (see **Table A1.1**).

Individual effects (α_i) are estimated as fixed effects. Panel data makes it possible to incorporate such effects because we observe each country in more than one time period, and they can be assumed to represent some unobservable cultural factors that are relatively constant over time.

Table A1.1. Variables used in the analysis

Name of the variable	Description of the variable	Source(s)
Explained (dependent) variable		
CASH_M1	The share of the currency in circulation in the sum of the currency in circulation and demand deposits held in financial institutions (M1 monetary aggregate), % of M1, seasonally adjusted. Minor deviations from this definition are possible for some countries. M1 monetary aggregate ("total transactional money") typically includes cash and demand deposits of residents (individuals and firms, excluding public administration), denominated in the national currency, held in local financial institutions. However, in some countries, the definition of M1 monetary aggregate deviates slightly from this standard definition and might include: foreign currency deposits, time deposits, deposits of the central governments and local (e.g., regional) governments. Exclusion of these additional components for all the countries that enter the econometric sample is impossible due to the lack of data on these non-standard components.	International Monetary Fund, Asian Development Bank, local central banks, own calculations
I. "Typical" cash shadow economy determinants ($\mathbf{x}_1$)		
TAX_TIME	Time to prepare and pay taxes in hours**	World Bank – Doing Business project
GCI_ETHICS	An indicator summarizing the ethical behavior of firms, based on answers to the following question: *"In your*	World Economic Forum – The Global Competitiveness Index

[12] The selection of countries took into account: (i) the necessity to include at least a couple of countries from each major region included in the Reports (Middle East and Africa, Latin America and Caribbean, Sub-Saharan Africa, Asia and Pacific), (ii) data availability and (iii) data quality. Not all the countries included in the Reports are covered by the econometric sample, but the variable coefficients can be used out-of-sample in order to estimate the shadow economy for each country discussed in the Reports.

	*country, how do you rate the corporate ethics of companies (ethical behavior in interactions with public officials, politicians and other firms)? [1 = extremely poor—among the worst in the world; 7 = excellent—among the best in the world]"***	
GCI_TRUST	An indicator summarizing the public trust in politicians, based on answers to the following question: *"In your country, how do you rate the ethical standards of politicians? [1 = extremely low; 7 = extremely high]"***	World Economic Forum – The Global Competitiveness Index
GCI_REG_BURDEN	An indicator summarizing the burden of government regulations, based on answers to the following question: *"In your country, how burdensome is it for companies to comply with public administration's requirements (e.g., permits, regulations, reporting)? [1 = extremely burdensome; 7 = not burdensome at all]"***	World Economic Forum – The Global Competitiveness Index
GCI_TRANSPARENCY	An indicator summarizing the transparency of government policymaking, based on answers to the following question: *"In your country, how easy is it for companies to obtain information about changes in government policies and regulations affecting their activities? [1 = extremely difficult; 7 = extremely easy]"***	World Economic Forum – The Global Competitiveness Index
GCI_CRIME_COST	An indicator summarizing the perceived business costs of crime and violence, based on answers to the following question: *"In your country, to what extent does the incidence of crime and violence impose costs on businesses? [1 = to a great extent—imposes huge costs; 7 = not at all—imposes no costs]"***	World Economic Forum – The Global Competitiveness Index
GCI_ORG_CRIME	An indicator summarizing the perceptions of organized crime, based on answers to the following question: *"In your country, to what extent does organized crime (mafia-oriented racketeering, extortion) impose costs on businesses? [1 = to a great extent—imposes huge costs; 7 = not at all—imposes no costs]"***	World Economic Forum – The Global Competitiveness Index
RULE_OF_LAW	The value of the indicator measuring the rule of law from the Worldwide Governance Indicators; ranges from approximately -2.5 (weak rule of law) to 2.5 (strong rule of law)**	World Bank – Worldwide Governance Indicators
GOV_EFFECTIVENESS	The value of the indicator measuring the government effectiveness from the Worldwide Governance Indicators; ranges from approximately -2.5 (low government effectiveness) to 2.5 (high government effectiveness)**	World Bank – Worldwide Governance Indicators
NON_EMPLOYED	The % share of unemployed (aged 15+*) and economically inactive (aged 15-64) persons in the total population (aged 15-64)**	International Labour Organization, own calculations
UNEMP	Unemployment rate, % of total labour force (economically active population), seasonally adjusted***	International Monetary Fund, World Bank
SELF_EMPLOYED	The ratio of the total number of self-employed (employers, own-account workers, members of producers' cooperatives, and contributing family workers) to the population aged 15-64**	International Labour Organization, own calculations
FAMILY_WORK	The ratio of the total number of contributing family workers to the population aged 15-64**	International Labour Organization, own calculations
OWN_ACCOUNT_WORK	The ratio of the total number of own-account workers to the population aged 15-64**	International Labour Organization, own calculations
II. Payment card system variables (x_2)		
CARDS_ACTIVE_PER_CAPITA	The number of active payment cards *per capita*, seasonally adjusted***. The number of active cards is calculated on the basis of additional data from the Global Findex database on the share of people using cards throughout the past year. We use the number of active payment cards *per capita* instead of payment cards transaction value, because the former variable is more likely to be exogenous (i.e. there are less feedback effects from the explained variable to the explanatory variables; such feedback effects are detrimental to the quality of estimation). In addition, we focus on active cards because these are more closely	Eurostat, European Central Bank, World Bank (Global Payment Systems Survey, Global Findex database), International Bank for Settlements, national central banks

	related to actual electronic payments – in fact, in some countries, the percentage of payments cards that are not used regularly is considerable	
III. Other control variables (x_3)		
GDP_REAL_PER_CAPITA	Real GDP *per capita* in PPS in constant 2010 prices (purchasing power parity adjusted, US dollar in 2011), seasonally adjusted***	Eurostat, International Monetary Fund, World Bank, own calculations
DEPOSIT_INTEREST_RATE	Households deposit rate *per annum*, in %***	International Monetary Fund, local central banks, for some countries simplified assumptions are made in order to translate interbank offer rates or central bank policy rates into household deposit rates
R_DEPOSIT_INTEREST_RATE	Households deposit rate per annum adjusted by yoy CPI inflation rate, in %***	International Monetary Fund, local central banks, for some countries simplified assumptions are made in order to translate the interbank offer rates or central bank policy rates into household deposit rates, own calculations
CPI_RATE	yoy CPI inflation rate, in %	Eurostat, International Monetary Fund, local central banks and statistical offices, for Argentina GDP deflator is used
CREDIT_GDP	Domestic credit to private sector [% of GDP], seasonally adjusted***	International Monetary Fund, World Bank, Bank for International Settlements, African Development Bank, local central banks and statistical offices
FIN_DEV	Index [0-1, 1=best development], aggregate of financial institutions, financial depth and financial market development indices**. IMF data available till 2014, for the 2015-2016 period, a forecast based on the **CREDIT_GDP** variable (see above) is formulated	International Monetary Fund – Svirydzenka (2016), own calculations based on the **CREDIT_GDP** variable (see above)
URBAN_POPULATION	The % share of urban population in the entire population**	World Bank
AGRI	The % share of people employed in agriculture in the overall employment**	International Labour Organization
INTERNET_ACCESS	The share of the population with Internet access, % of population**	International Telecommunication Union (United Nations)
DUMMY_IND	Binary variable controlling for the effect of demonetization in India in Q4 2016	Own elaboration
DUMMY_ROU	Binary variable controlling for the credit boom in Romania starting in Q1 2007	Own elaboration

* - Data for unemployed aged 15-64 is unavailable, but unemployed persons are most likely less than 65 years old.
** - interpolated from annual to quarterly.
*** - for some countries interpolated from annual to quarterly.
Source: EY elaboration. The historical data on Global Competitiveness Index was provided by the courtesy of World Economic Forum representatives.

Table A1.1 contains many variables which cannot be included in a single econometric specification at once due to sample size constraints and collinearity. In order to formulate the best CDA model, we apply a formal variable selection procedure. The selection is based on the following principle: the final set of explanatory variables $\mathbf{x_1}$, $\mathbf{x_2}$ and $\mathbf{x_3}$ should produce a possibly parsimonious model, but also achieve a satisfactory fit of the dependent variable.

The assessment is based on the model averaging procedure in which a wide array of variants of equation (A1.1) is estimated using the Panel Corrected Standard Errors (PCSE) method[13], with different combinations of variables from **Table A1.1.** The set of econometric specifications for the procedure is constructed so as to:

- always include the **GDP_REAL_PER_CAPITA**, **DUMMY_IND**, **DUMMY_ROU** variables and fixed effects (α_i). These variables, belonging to the other control variables group ($\mathbf{x_3}$), play a vital role in explaining the structural demand for cash;
- include at least one cash shadow economy determinant from the group $\mathbf{x_1}$ so that the CDA model can effectively serve as a tool for decomposition of cash demand;
- avoid including variables with similar information content. From the following clusters of variables, only a single one could be included in a given econometric specification:
    - 1: Global Competitiveness Index variables [**GCI_TRUST, GCI_REG_BURDEN, GCI_TRANSPARENCY, GCI_CRIME_COST, GCI_ORG_CRIME, GCI_ETHICS**],
    - 2: forms of self-employment [**FAMILY_WORK, SELF_EMPLOYED, OWN_ACCOUNT_WORK**],
    - 3: interest rate [**DEPOSIT_INTEREST_RATE**, **R_DEPOSIT_INTEREST_RATE**],
    - 4: non-activity in the labour market [**NON_EMPLOYED**, **UNEMP**],
    - 5: attitude to government and law [**RULE_OF_LAW**, **GOV_EFFECTIVENESS**],
    - 6: financial variables [**FIN_DEV**, **CREDIT_GDP**].

To make the number of models to estimate manageable (each additional variable can multiply the total number of models by a factor of two), in this step no interaction terms are considered (the only exception being the square of **GDP_REAL_PER_CAPITA** included in order to account for non-linearity). In total, S=241,440 models meet all the criteria specified above. Note that there are different data points missing for different variables, so that a given selection of variables implies the maximum sample size. To avoid the differences between the models resulting from this implicit sample selection, the same minimum sample of 1812 is used (for 64 countries), common to all the models under consideration. In each case, all the variables (including the **CASH_M1** variable) are transformed using the Prais-Winsten transformation.

In order to summarise the estimation results and select appropriate variables to the final model, two kinds of weighted average results are calculated:

- The marginal probability of a given variable being included in the "true" model,
- The weighted average value of particular coefficients.

Each estimation result is weighted according to formal criteria describing the quality of each specification – the "better" the specification, the higher the weight. For robustness, we considered two alternative weighting schemes:

- **The frequentist weight** based on Akaike Information Criterion (AIC), calculated according to the following formula[14]:

$$\boldsymbol{w_s^{Freq}} = \frac{\boldsymbol{exp(-\frac{AIC_s - \min\limits_s(AIC_s)}{2})}}{\boldsymbol{\sum_{s=1}^{S} exp(-\frac{AIC_s - \min\limits_s(AIC_s)}{2})}}, \text{ (A1.3)}$$

$$\boldsymbol{AIC_s = N log\left(\frac{SSR_s}{N}\right) + 2K_s}, \text{ (A1.4)}$$

in which $\mathbf{s=1,2,\ldots,}\boldsymbol{S}$ indexes the models under consideration; $\boldsymbol{N}$ is the total number of observations; $\boldsymbol{SSR_s}$ is the sum of squared, Prais-Winsten-transformed residuals; and $\boldsymbol{K_s}$ is the number of

---

[13] The method is robust to: contemporaneous correlation of error terms between panel units, serial correlation of order 1 of the error term (a common serial correlation coefficients for all the panels is selected) as well as to heteroskedasticity.

[14] The formula (A1.3) is based on Buckland S. T., Burnham K. P. and Augustin N. H. (1997), "Model Selection: An Integral Part of Inference", *Biometrics*, Vol. 53(2), pp. 603-618.

coefficients that are not automatically included in all the specifications, i.e., it is the total number of estimated coefficients minus 67 (**GDP_REAL_PER_CAPITA**, **DUMMY_IND**, **DUMMY_ROU**, 64 country dummies). In other words, the AIC information criterion grows when a given model performs worse in terms of fitting the data ($\boldsymbol{SSR_s}$), and additionally penalises specifications that use too many variables ($\boldsymbol{K_s}$). Models with a lower value of AIC are thus preferred and weighted higher.

- **The Bayesian weight** based on the posterior probability that a given specification is "true". The respective calculations are part of Bayesian statistical methods that combine the expert knowledge (prior information) with the information provided by empirical data. The weight is calculated according to the following formula[15]:

$$\boldsymbol{w_s^{Bayes}} = \frac{\boldsymbol{\Pi_s^{\mathbf{posterior}}}}{\boldsymbol{\sum_{s=1}^{S} \Pi_s^{\mathbf{posterior}}}}, \text{ (A1.5)}$$

$$\boldsymbol{\Pi_s^{\mathbf{posterior}}} = \boldsymbol{\Pi_s^{\mathbf{prior}} exp\left(\lambda_s - \max_s(\lambda_s)\right)}, \text{ (A1.6)}^{16}$$

$$\boldsymbol{\lambda_s} = -\frac{\boldsymbol{K_s}}{\mathbf{2}} \boldsymbol{ln(1+g)} - \frac{\boldsymbol{N-1}}{\mathbf{2}} \boldsymbol{ln(\frac{1}{1+g} + \frac{g}{1+g}(1 - R_s^2))}, \text{ (A1.7)}$$

$$\boldsymbol{\Pi_s^{\mathbf{prior}}} = \boldsymbol{\theta^{K_s}(1-\theta)^{K_{max}-K_s}}, \text{ (A1.8)}$$

$$\boldsymbol{\theta} = \frac{\boldsymbol{K_{max}-1}}{\boldsymbol{K_{max}}}, \text{ (A1.9)}$$

$$\boldsymbol{K_{max}} = \boldsymbol{\max_s(K_s)}, \text{ (A1.10)}$$

$$\boldsymbol{R_s^2} = \boldsymbol{1 - \frac{SSR_s}{SST}}, \text{ (A1.11)}$$

$$\boldsymbol{g} = \boldsymbol{(K_{max})^2}, \text{ (A1.12)}$$

in which $\boldsymbol{SST}$ is the variance of Prais-Winsten-transformed **CASH_M1** variable (the explained variable); $\boldsymbol{\Pi_s^{\mathbf{prior}}}$ is the prior probability that the model is true (it summarises the prior expectations with respect to, among other things, the number of variables in the true model); $\boldsymbol{R_s^2}$ is the coefficient of determination summarising the quality of model fit (the larger the better); and the $\boldsymbol{K_{max}}$ parameter equals 13 (i.e. model with variables representing all 6 variable clusters and 7 additional variables – see Table A1.3).

Both in the frequentist and in the Bayesian case, an additional variant of model averaging is considered, in which specifications with economically unexpected signs of the variables belonging to $\mathbf{x_1}$ and $\mathbf{x_2}$ are removed. Such sign restrictions are described in **Table A1.2.** and this version of weights is further referred to as *restricted* (as opposed to *unrestricted*).

[15] The formulas are based on Steel M. F. J. (2017), "Model Averaging and its Use in Economics", MPRA Paper No. 81568.

[16] The original versions of the formulae (A1.3) and (A1.6) are modified by subtracting $\boldsymbol{\max_s(AIC_s)}$ and $\boldsymbol{\max_s(\lambda_s)}$ from $\boldsymbol{AIC_s}$ and $\boldsymbol{\lambda_s}$, respectively. Note that this is arithmetically neutral for the computation of $\boldsymbol{w_s^{Freq}}$ and $\boldsymbol{w_s^{Bayes}}$ in (A1.3) and (A1.5). This modification, however, helps to avoid numerical problems.

Table A1.2. The restrictions on the sign of the coefficients in each specification entering the restricted model averaging procedure

Name of the variable	The direction of impact on cash demand and shadow economy*
TAX_TIME	**+**
GCI_ETHICS	**-**
GCI_TRUST	**-**
GCI_REG_BURDEN	**-**
GCI_TRANSPARENCY	**-**
GCI_CRIME_COST	**-**
GCI_ORG_CRIME	**-**
RULE_OF_LAW	**-**
GOV_EFFECTIVENESS	**-**
NON_EMPLOYED	**+**
UNEMP	**+**
SELF_EMPLOYED	**+**
FAMILY_WORK	**+**
OWN_ACCOUNT_WORK	**+**
CARDS_ACTIVE_PER_CAPITA	**-**

*The impact is allowed to be zero in each case.
Source: EY elaboration.

The probabilities of inclusion of particular variables in the specification are presented in **Table A1.3**. In each case, we consider frequentist and Bayesian weights with restricted and non-restricted coefficient signs. The restricted variants do not differ substantially from the unrestricted variant when it comes to the probability of inclusion. The variables in bold (e.g., **GCI_ETHICS, RULE_OF_LAW**) are those variables for which the respective probabilities are larger than 30% in any column and which are thus the main candidates for inclusion in the final model. However, in order to verify that those variables should be included, the average coefficients need to be analyzed as well – they are presented in **Table A1.4**.

Table A1.3. The probability of inclusion of particular variables in the final specification under different model averaging assumptions

Cluster		Frequentist weights		Bayesian weights	
		Unrestricted	*Restricted*	*Unrestricted*	*Restricted*
	I. "Typical" cash shadow economy determinants (x_1)				
No	**TAX_TIME**	**88.8%**	**87.5%**	**93.1%**	**91.6%**
1	**GCI_ETHICS**	**63.1%**	**66.8%**	**58.8%**	**66.4%**
1	GCI_TRUST	3.8%	3.4%	5.3%	4.3%
1	GCI_REG_BURDEN	4.6%	4.1%	6.2%	4.9%
1	GCI_TRANSPARENCY	3.7%	3.4%	5.2%	4.3%
1	GCI_CRIME_COST	9.7%	8.8%	11.9%	9.8%
1	GCI_ORG_CRIME	5.8%	5.1%	7.5%	6.0%
5	**RULE_OF_LAW**	**92.3%**	**92.8%**	**91.2%**	**92.2%**
5	GOV_EFFECTIVENESS	4.9%	4.7%	7.0%	6.3%
4	**NON_EMPLOYED**	**98.7%**	**99.1%**	**97.5%**	**98.3%**
4	UNEMP	0.7%	0.6%	1.7%	1.3%
2	SELF_EMPLOYED	16.7%	6.4%	22.9%	13.6%
2	**FAMILY_WORK**	**18.1%**	**27.2%**	**25.0%**	**43.6%**
2	OWN_ACCOUNT_WORK	21.0%	0.0%	27.7%	0.0%
	II. Payment card system variables (x_2)				
	CARDS_ACTIVE_PER_CAPITA	**27.7%**	**19.3%**	**48.7%**	**37.8%**

	III. Other control variables (x_3)				
-	**GDP_REAL_PER_CAPITA**	**Always included**	**Always included**	**Always included**	**Always included**
No	**GDP_REAL_PER_CAPITA^2**	**32.2%**	**31.9%**	**53.3%**	**51.8%**
3	**DEPOSIT_INTEREST_RATE**	**100.0%**	**100.0%**	**100.0%**	**100.0%**
3	R_DEPOSIT_INTEREST_RATE	0.0%	0.0%	0.0%	0.0%
No	**CPI_RATE**	**61.0%**	**61.9%**	**76.0%**	**77.2%**
6	**CREDIT_GDP**	**31.2%**	**34.2%**	**41.7%**	**48.1%**
6	**FIN_DEV**	**20.9%**	**19.4%**	**30.2%**	**24.9%**
No	**URBAN_POPULATION**	**50.4%**	**53.8%**	**68.2%**	**72.1%**
No	**AGRI**	**28.3%**	**31.1%**	**49.2%**	**54.5%**
No	**INTERNET_ACCESS**	**94.8%**	**93.1%**	**95.9%**	**94.1%**
-	**DUMMY_IND**	**Always included**	**Always included**	**Always included**	**Always included**
-	**DUMMY_ROU**	**Always included**	**Always included**	**Always included**	**Always included**
	Additional information				
	Number of models	241440	106137	241440	106137
	The avg. number of coefficients	75.79	75.55	77.24	77.03

Notes: fixed effects α_i omitted for brevity of presentation, but included in each specification, as well as **GDP_REAL_PER_CAPITA**, **DUMMY_IND**, **DUMMY_ROU** variables. The variables in bold are those variables for which the probabilities of inclusion in the "true" model are larger than 30% in any case and which are considered for inclusion in the model. The number of observations equals 1812, the number of included countries equals 64 and the number of coefficients ranges from 68 to 80. Each model was estimated using the PCSE method with common serial correlation coefficient of the error term calibrated at the level of 0.83 (based on our preliminary analyses). The restricted variants require that each specification meets the sign restrictions as enumerated in **Table A1.2.**
Source: EY elaboration.

Once again, the main candidates for inclusion in the final specification are indicated in bold. Note that the results are quite robust to imposition of sign restrictions (with some minor exceptions among variables bearing low inclusion probabilities). Only in the case of three variables out of those that are indicated in bold has the corresponding coefficient counter-intuitive sign: namely, **CREDIT_GDP** and **URBAN_POPULATION** have a positive impact, while **AGRI** has a negative impact on the demand for cash in all four weighting schemes. We mark these variables in red in **Table A1.4**.

Table A1.4. The weighted average coefficients in the considered econometric models under different model averaging assumptions

	Frequentist weights		Bayesian weights	
	Non-restricted	*Restricted*	*Non-restricted*	*Restricted*
I. "Typical" cash shadow economy determinants (x_1)				
TAX_TIME	0.004	0.004	0.004	0.004
GCI_ETHICS	-0.541	-0.573	-0.504	-0.570
GCI_TRUST	-0.005	-0.005	-0.007	-0.006
GCI_REG_BURDEN	-0.011	-0.010	-0.016	-0.012
GCI_TRANSPARENCY	-0.005	-0.004	-0.007	-0.005
GCI_CRIME_COST	-0.034	-0.030	-0.042	-0.034
GCI_ORG_CRIME	-0.014	-0.012	-0.018	-0.014
RULE_OF_LAW	-2.416	-2.433	-2.406	-2.434
GOV_EFFECTIVENESS	-0.067	-0.062	-0.095	-0.083
NON_EMPLOYED	0.257	0.262	0.248	0.256
UNEMP	0.001	0.001	0.002	0.001
SELF_EMPLOYED	-0.001	0.001	-0.001	0.001
FAMILY_WORK	0.008	0.013	0.013	0.022
OWN_ACCOUNT_WORK	-0.014	0.000	-0.018	0.000
II. Payment card system variable (x_2)				
CARDS_ACTIVE_PER_CAPITA	-0.023	-0.036	-0.041	-0.066
III. Other control variables (x_3)				
GDP_REAL_PER_CAPITA	-0.050	-0.051	-0.155	-0.153
GDP_REAL_PER_CAPITA^2	0.006	0.006	0.009	0.009
DEPOSIT_INTEREST_RATE	0.153	0.153	0.155	0.155
R_DEPOSIT_INTEREST_RATE	0.000	0.000	0.000	0.000
CPI_RATE	-0.014	-0.014	-0.018	-0.018
CREDIT_GDP	0.003	0.004	0.004	0.005
FIN_DEV	-0.262	-0.238	-0.385	-0.308
URBAN_POPULATION	0.078	0.084	0.105	0.111
AGRI	-0.005	-0.007	-0.008	-0.012
INTERNET_ACCESS	-0.035	-0.035	-0.036	-0.036
DUMMY_IND	-20.163	-20.162	-20.170	-20.165
DUMMY_ROU	-16.558	-16.565	-16.552	-16.563
Additional information				
Number of models	241440	106137	241440	106137
The avg. number of coefficients	75.79	75.55	77.24	77.03

Notes: fixed effects, α_i, are omitted for brevity of presentation, but included in each specification, as well as **GDP_REAL_PER_CAPITA**, **DUMMY_IND**, **DUMMY_ROU** variables. The variables in bold are those variables for which the probabilities of inclusion in the "true" model are larger than 30% for at least one of the weighting schemes and which are considered for inclusion in the model. The rows in red indicate variables with counter-intuitive sign of the corresponding coefficients. The number of observations equals 1812, the number of included countries equals 64 and the number of coefficients ranges from 68 to 80. Each model was estimated using the PCSE method with common serial correlation coefficient of the error term calibrated at the level of 0.83 (based on our preliminary analyses). The restricted variants require that each specification meets the sign restrictions as outlined in the **Table A1.2.**
Source: EY elaboration.

The remaining candidate coefficients have appropriate signs and are indicated in green with the following exceptions: **TAX_TIME** and **INTERNET_ACCESS**. Inclusions of these variables in the final models means that other, economically crucial variables - **CARDS_ACTIVE_PER_CAPITA** and **GDP_REAL_PER_CAPITA** - become statistically insignificant due to collinearity issues (internet access and prevalence of electronic cards are clearly related with GDP level). This is why we do not include **TAX_TIME** and **INTERNET_ACCESS** in the final model and focus only on the variables marked in green in **Table A1.4**.

As a further step, we verify which interaction terms should be included in the model (as those are not subject to the model averaging procedure). Here we apply a general-to-specific procedure with a starting point being a "Full model" in which the interaction terms are applied to all the $\mathbf{x_1}$ and $\mathbf{x_2}$ variables (as grouped in **Table A1.1**). The only interaction among other control variables ($\mathbf{x_3}$) is the square of the **GDP_REAL_PER_CAPITA**, which is considered an obligatory variable in order to better account for possible non-linearities related to socio-economic development of countries included in the sample. The relevance of this variable was confirmed in the model averaging procedure. Note that we do not apply interactions to all the variables belonging to $\mathbf{x_3}$ in order to keep each specification as parsimonious as possible.

The starting and final points of the general-to-specific procedure are demonstrated in **Table A.1.5**. We sequentially eliminated the interaction terms with highest p-values.[17] We treat the results in column (2) as the baseline model. There are two statistically insignificant variables remaining (**CPI_RATE** and **FIN_DEV**), but they play an important economic role and remain included as a result of the model averaging procedure.

The estimates of the overall level of the shadow economy, based on the described methodology, and the results included in column (4) of **Table A1.5** are discussed in the Reports. While calculating the shadow economy we assume that the coefficient that depends on **GDP_REAL_PER_CAPITA** cannot violate the sign restrictions included in **Table A.1.2** – if the calculated coefficient has inappropriate sign, we assume that it is equal to zero in a particular period for a given country**. This correction becomes effective in the case of FAMILY_WORK** interaction variable (in wealthier countries), and – to a much lesser extent – the **RULE_OF_LAW** interaction.

[17] We also eliminated the interaction term between **GDP_REAL_PER_CAPITA** and **CARDS_ACTIVE_PER_CAPITA** because its sign was economically unintuitive.

Table A1.5. Econometric estimates of the currency demand equation in the general-to-specific procedure

	(1) Full model	(2) Baseline (reduced) model
GCI_ETHICS	-1.256* (0.062)	-1.047*** (0.006)
GDP_REAL_PER_CAPITA ×GCI_ETHICS	0.0576 (0.511)	
RULE_OF_LAW	-4.116*** (0.003)	-3.820*** (0.002)
GDP_REAL_PER_CAPITA × RULE_OF_LAW	0.355** (0.037)	0.268* (0.065)
NON_EMPLOYED	0.214 (0.120)	0.213** (0.016)
GDP_REAL_PER_CAPITA ×NON_EMPLOYED	0.00255 (0.894)	
FAMILY_WORK	0.358** (0.037)	0.324* (0.054)
GDP_REAL_PER_CAPITA × FAMILY_WORK	-0.208*** (0.008)	-0.194** (0.012)
CARDS_ACTIVE_PER_CAPITA	0.701 (0.547)	-1.228** (0.012)
GDP_REAL_PER_CAPITA ×CARDS_ACTIVE_PER_CAPITA	-0.274** (0.020)	
GDP_REAL_PER_CAPITA	-1.557 (0.241)	-1.285* (0.091)
GDP_REAL_PER_CAPITA ×GDP_REAL_PER_CAPITA	0.0465** (0.038)	0.0416* (0.070)
DEPOSIT_INTEREST_RATE	0.160*** (0.000)	0.168*** (0.000)
CPI_RATE	-0.0243 (0.187)	-0.0236 (0.200)
FIN_DEV	-1.660 (0.412)	-1.305 (0.512)
DUMMY_IND	-20.10*** (0.000)	-20.02*** (0.000)
DUMMY_ROU	-16.42*** (0.000)	-16.37*** (0.000)
The number of observations	1881	1881
The number of countries	64	64
The AR(1) coefficient in Prais-Winsten transformation	0.836	0.834

Notes: p-values in parentheses, * p<0.1, ** p<0.05, *** p<0.01. 64 α_i dummies are omitted for brevity of presentation, but included in each specification. All the models are estimated using the PCSE estimator with common AR component of the error term.
Source: EY elaboration.

Substep 3.2. Using CDA model to measure the shadow-economy-related cash

In the second substep, we set the values of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ vectors in equation (1) at their "best" observable levels for the countries included in the sample, e.g.: the lowest value of **NON_EMPLOYED**, the highest number of payment cards, etc., and estimate the theoretical value of the explained variable that would prevail in the case of the lowest possible cash shadow economy.

Importantly, the aforementioned theoretical value of the dependent variable is computed from equation (A1.1) using the estimated value of $\boldsymbol{\beta}_{i,t}^{(1)}$ and only a fraction of $\boldsymbol{\beta}_{i,t}^{(2)}$ (further referred to as $\boldsymbol{\beta}_{i,t}^{(2,1)}$). This is because we do not treat the payment card system variable ($\mathbf{x_2}$) as a typical shadow economy determinant, since it also reduces the structural demand for cash. Hence, a symmetric treatment of the coefficients $\boldsymbol{\beta}_{i,t}^{(1)}$ and $\boldsymbol{\beta}_{i,t}^{(2)}$ would most likely result in an overestimation of the shadow economy size and, consequently, of the positive effects of promoting cashless payments. Instead, we split the estimated coefficient $\boldsymbol{\beta}_{i,t}^{(2)}$ into two components – related and unrelated to the shadow economy, by assuming that the proportion of their effects on the dependent variable is analogous to the proportion of (i) the average impact of the typical shadow economy determinants ($\mathbf{x_1}$) on the level of the dependent variable and (ii) an analogous value for other control variables ($\mathbf{x_3}$). In other words, we split $\boldsymbol{\beta}_{i,t}^{(2)}$ proportionally to the variability of the dependent variable explained by $\mathbf{x_1}$ and $\mathbf{x_3}$, say $\boldsymbol{\beta}_{i,t}^{(2,1)}$ and $\boldsymbol{\beta}_{i,t}^{(2,3)}$ ($\boldsymbol{\beta}_{i,t}^{(2)} = \boldsymbol{\beta}_{i,t}^{(2,1)} + \boldsymbol{\beta}_{i,t}^{(2,3)}$). In this calculation, we include the interaction terms (classified as in Table A1.5), but exclude the dummy variables (fixed effects, **DUMMY_IND** and **DUMMY_ROU**).

The difference between the fitted value from the model, calculated on the basis of the factual values of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and the counterfactual theoretical value may be interpreted as the share of cash in the M1 aggregate that is related to shadow economy monetary transactions ($\frac{\boldsymbol{C}_{i,t}^{SHADOW}}{\boldsymbol{M1}_{i,t}}$). Given the observed stock of the M1 aggregate for a given country and period, the obtained difference allows us to calculate the amount of cash that is attributable to the cash shadow economy ($\boldsymbol{C}_{i,t}^{SHADOW}$).

Substep 3.3. Conversion of the shadow cash estimate into the cash shadow economy estimate

In the third substep, we estimate the size of the cash shadow economy[18] on the basis of the quantity theory of money and using additional information about the total economy and the monetary economy obtained in the Steps 1 and 2. First, we assume that the *velocity* of money in the cash shadow economy is equal to the velocity of money in the overall monetary economy:

$$\frac{\boldsymbol{Y}_{i,t}^{MONETARY,TOTAL}}{\boldsymbol{M1}_{i,t}} = \frac{\boldsymbol{Y}_{i,t}^{CASH,SHADOW}}{\boldsymbol{C}_{i,t}^{SHADOW}}, \text{ (A1.13)}$$

where $\boldsymbol{Y}_{i,t}^{MONETARY,TOTAL}$ and $\boldsymbol{Y}_{i,t}^{CASH,SHADOW}$ denote the monetary output in the total and shadow economy, respectively; $\boldsymbol{C}_{i,t}^{SHADOW}$ stands for the amount of cash used for settling transactions in the cash shadow economy and $\boldsymbol{M1}_{i,t}$ is the M1 total transactional money.[19]

We transform equation (A1.13) to estimate the share of the cash shadow economy output in the total monetary output (including also the cash shadow economy) without knowing the exact value of the velocity of money*:*

$$\frac{\boldsymbol{Y}_{i,t}^{CASH,SHADOW}}{\boldsymbol{Y}_{i,t}^{MONETARY,TOTAL}} = \frac{\boldsymbol{C}_{i,t}^{SHADOW}}{\boldsymbol{M1}_{i,t}}. \text{ (A1.14)}$$

[18] The size of the cash shadow economy corresponds to the part of monetary output / monetary GDP that is generated in the shadow economy.

[19] The assumption of equal velocity in the overall and shadow monetary economy may not always be fulfilled. E.g., in South Africa, the velocity of notes and coins is higher than the velocity of the remaining part of M1 (i.e., the demand deposits). In other cases, however, demand deposits might be transferred around the country much faster than cash. The violation of the equal velocity assumption might also be related to dollarization in various ways. In such a case, the ratio **CASH_M1** (see Table A1.1) does not correspond to total demand for cash. All in all, we assume that the equal velocity assumption is fulfilled on average so that the potential error averages out across countries in the panel.

Note that $\frac{C_{i,t}^{SHADOW}}{M1_{i,t}}$ is the endpoint of the substep 3.2. However, it is only related to those economic activities that include monetary transactions. In order to obtain the estimate of the total shadow economy $\boldsymbol{Y_{i,t}^{TOTAL,SHADOW}}$ (as a share in total economy $\boldsymbol{Y_{i,t}^{TOTAL}}$), we use the following formula:

$$\frac{Y_{i,t}^{TOTAL,SHADOW}}{Y_{i,t}^{TOTAL}} = \frac{Y_{i,t}^{CASH,SHADOW}}{Y_{i,t}^{MONETARY,TOTAL}} \times \frac{Y_{i,t}^{MONETARY,TOTAL}}{Y_{i,t}^{TOTAL}} + \frac{Y_{i,t}^{NMSE}}{Y_{i,t}^{TOTAL}}, \text{(A1.15)}$$

in which $\frac{Y_{i,t}^{MONETARY,TOTAL}}{Y_{i,t}^{TOTAL}}$ is the output of Step 2 and the $Y_{i,t}^{NMSE}$ is the non-monetary shadow economy estimated separately (see Step 2 for details). Finally, the share of the total shadow economy in the official GDP estimate is obtained using the following adjustment:

$$\frac{Y_{i,t}^{TOTAL,SHADOW}}{Y_{i,t}^{OFFICIAL}} = \frac{Y_{i,t}^{TOTAL,SHADOW}}{Y_{i,t}^{TOTAL}} \times \frac{Y_{i,t}^{TOTAL}}{Y_{i,t}^{OFFICIAL}}, \text{(A1.16)}$$

in which $\frac{Y_{i,t}^{OFFICIAL}}{Y_{i,t}^{TOTAL}}$ is the result of the Step 1.

The approach presented here is superior to the approach often applied in the literature, consisting in translating the value of cash used for settling shadow economy transactions (obtained from the currency demand analysis) into the value of the shadow economy output with the use of some specific estimate of the velocity of money (sometimes based on dubious and/or outdated figures). The problem arises because, in order to calculate the velocity of money, one should divide either (i) the total GDP ($\boldsymbol{Y_{i,t}^{TOTAL}}$, including the best available estimate of shadow economy) by total M1 or (ii) GDP that excludes any estimate of the shadow economy ($\boldsymbol{Y_{i,t}^{TOTAL} - Y_{i,t}^{TOTAL,SHADOW}}$) by the estimated value of M1 that is not related to shadow economy activities. In practice, many researchers simply divide the official GDP figure (as published by the statistical offices) by total M1 assuming implicitly that there is no shadow economy in the official GDP estimate.[20] This approach is controversial for two reasons. Firstly, if official GDP figures do not include any shadow economy estimate, this leads to an underestimation of the velocity (since the part of GDP that is related to the shadow economy is missing in the numerator of the ratio, but the shadow economy related to cash is included in the denominator – M1). Secondly, if some shadow economy estimate is already included in the official GDP, then the above-mentioned calculation of the money velocity already assumes the validity of one shadow economy estimate. When another shadow economy estimate is obtained using this velocity as an input, the method could be regarded as intrinsically inconsistent.

Another solution, described by Ardizzi et al.[21], Warner et. al.[22] and mentioned by Schneider et al.[23], is to assume that the share of the shadow economy in the official GDP in some base period equals zero, and the velocity of money is obtained by dividing the value of GDP in that period by the "legal" component of M1 (obtained after subtracting the "shadow" money value that was estimated with the use of the CDA). Assuming that the velocity of money is constant over time, one may then calculate the monetary level of the shadow economy in all the periods of the analysis. However, we find the constant velocity assumption highly implausible.

Step 4. Measuring the passive shadow economy by using labour market analysis

Passive shadow economy consists in underreporting of the revenues by registered, legally operating entities. We assume that the remaining part of the shadow economy, i.e. the committed shadow

[20] Cf. Tanzi (1983), *op. cit.*, Embaye (2007), *op. cit.*, Pickhardt, M., Sarda, J. (2010), "The size of the underground economy in Germany: A correction of the record and new evidence from the Modified-Cash-Deposit-Ratio approach", CAWM discussion paper, Centrum für Angewandte Wirtschaftsforschung Münster, No. 36.

[21] Ardizzi et al. (2014), op cit.

[22] Kyle S. C., Warner, A., Dimitrov L., Krustev R., Alexandrova S., Stanchev K. (2001), "Measuring the Shadow Economy in Bulgaria," Working Papers 127656, Cornell University, Department of Applied Economics and Management.

[23] Schneider F., Buehn A., Montenegro C. E. (2010), "Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007," *Policy Research Working Paper* 5356, The World Bank.

economy and the non-monetary shadow economy, generates value added that is correlated with and mirrored by shadow labour force inputs. Consequently, we treat the informal employment share in total employment as a starting point in the approximation of the share of committed shadow economy and non-monetary shadow economy activities in the total shadow economy.[24]

To decompose the overall level of the shadow economy into different components, we collect data on the share of informal employment in the total employment. Following the International Labour Organization (ILO) definition, the informal employment covers unregistered employees and people working in unregistered enterprises (enterprises that are not constituted as legal entities).[25] For the highest accuracy of estimates, we use various sources of informal employment data. We consider estimates provided by statistical offices as the primary source; however, only some of the statistical offices provide such data (e.g. in Algeria, Colombia, Egypt, Kenya, Senegal and South Africa). Another source that we use (when no official estimates are available) are the estimates provided by ILO. In some countries, we estimate the share of informal employment in the total employment on the basis of a comparison of two different measures of employment: data on the officially registered employment and data from the Labour Force Surveys based on declarations in the survey questionnaires. The latter source aims at measuring the total employment, including activities in the shadow economy. In addition to this, we estimated a panel econometric model based on informal employment data collected for 48 countries (see Table A1.6) that allows us to impute the remaining missing data (e.g. if estimates are not provided on an annual basis).

Table A1.6. Supplementary econometric model of the share of informal employment in total employment

Variables	Coefficient (p-value)
The share of self-employed people in total employment	0.1605** (0.017)
Poverty rate (% of people living below the poverty threshold)	0.0462 (0.168)
The share of low-skill employees (% of employees performing low-skill jobs)	0.2560** (0.011)
Labour productivity (output per worker)	-0.1244** (0.015)

Notes: p-values in parentheses, * p<0.1, ** p<0.05, *** p<0.01.
Fixed effects estimation. Country dummies are omitted for brevity of presentation, but included in our specification.
Source: EY elaboration.

The share of informal employment in the total employment does not imply the same share of the sum of the committed and the non-monetary shadow economy in GDP. We account for two additional factors here. First, we take into consideration the fact that, on average, labour productivity of officially employed people is higher than of those working in the shadow economy. We assume that the value added generated by the shadow labour force is related to remunerations. Therefore, we multiply the share of informal employment in the total employment by the labour income share in GDP which yields the initial ratio of the committed shadow economy plus non-monetary shadow economy to the total GDP. Next, we multiply the obtained result by the ratio of the minimum wage to the average monthly wage (or the ratio of the average monthly wage paid for performing elementary occupations to the official average monthly wage if there is no minimum wage) to account for the lower level of remunerations of the informal workers.

Such an approach is essentially an approximation of the actual labour productivity in the informal employment. The estimates of the average remuneration in the shadow economy are rarely available and for most of the countries it remains unclear whether this remuneration is lower or higher than the official minimum wage. Some firms avoid registration to pay wages below the minimum level, which indicates that shadow economy wages can be lower than the minimum wage. This would imply

[24] It should be noted, however, that unreported employment is also possible in the registered entities that are not involved in the committed shadow economy and the non-monetary shadow economy. On the other hand, we do not account for the fact that some companies with no unreported labour force may also be involved in the committed shadow economy.

[25] http://www.ilo.org/ilostat-files/Documents/description_IFL_EN.pdf (accessed 17.07.2018)

that our approach to estimating the productivity of unregistered workers may result in overestimation of the committed and the respective underestimation of the passive component of the shadow economy. This, in turn, is consistent with our preference for conservative estimation. On the other hand, we assume that the whole production value is used to pay the remuneration of informal workers. This assumption should not be controversial in the case of own-account workers (constituting majority of informal workers), but in some unregistered companies an over-proportional part of the revenue may be earned by the owner of such an entity.

Second, dedicated research for some (developed) countries indicates that the number of hours worked by an officially employed person is, on average, higher than in the case of a person working in the shadow economy. To estimate this difference, we use the LFS data on working hours for both the official and the informal employment (we assume that average weekly time in elementary occupations provides a proxy for working time in informal jobs).

After applying the above-discussed corrections, the adjusted estimate of the share of informal employment in the total employment allows us to approximate the share of the committed shadow economy and the non-monetary shadow economy in the total economy, Expressing other estimates also as shares in the total economy, we subtract from the two components our estimate of the non-monetary shadow economy (calculated in step 2) to arrive at the committed shadow economy estimate.[26] Finally, to obtain the passive shadow economy estimate, we subtract the committed shadow economy from our cash shadow economy estimate (calculated in substep 3.3).

[26] In the countries where the value added of agriculture exceeds 10% of GDP, instead of subtracting only the non-monetary shadow economy estimate we subtract the larger of the two: either (1) the non-monetary shadow economy estimate or (2) our estimate of the value added of agriculture generated by informal labour (which should already include the non-monetary shadow economy, since non-monetary shadow economy activities are related mostly to agriculture). Countries where agriculture still plays an important role in the economy often do not require registration of small businesses operating in this sector. Therefore, in such countries, agricultural economic activity related to informal labour does not correspond with our definition of the committed shadow economy (see chapter 1.3 in the Reports).

Appendix 2. Estimation of the value of consumer cash payments

One side of the passive shadow economy transactions are consumers. Therefore, the passive shadow economy size is related to the value of consumer cash payments in the economy. While estimating the impact of various regulations on the change in the passive shadow economy, we assume that (i) the value share of cash consumer payments crowded out by non-cash payments is equal to (ii) the percentage decline in the passive shadow economy size[27] (see also Charts 4.1-4.2 in Chapter 4 of the Reports). The numerator of the share (i), i.e. the value of additional cashless payments generated by a given measure, is country and regulation specific – we describe our calculations in this area for different regulations in Appendix 3. The share's denominator, i.e. the value of consumer cash payments before introducing any regulation, is also country specific. In this appendix we present our approach to the estimation of this value for different countries. This process may be divided into the following steps:

1. We start with the value of **household final consumption expenditure (HFCE)**, a figure in the national accounts systems. In the case of most countries, this data is extracted from the World Bank's database. Otherwise, we refer to the national sources.
2. For the countries where the value of the shadow economy estimate included in GDP is lower (higher) than EY estimate, we add (subtract) the absolute difference between the two estimates to (from) HFCE. This step is based on the assumption that most unreported economic activity is related to final expenditure of consumers (not final expenditure of government, enterprises (investment) or entities abroad (exports)).[28] This way we obtain our best estimate of the total value of HFCE for a given economy (accounting for the size of the shadow economy).
3. From the value obtained in step 2 we subtract the estimates of two components of non-monetary consumption: (i) the value of imputed rents (see footnote 5) that are included in HFCE, obtained from the statistical offices; (ii) the value of household production of goods for own final use (the non-monetary shadow economy, see footnote 8) that is either included in original HFCE figures (when national account statistics include non-monetary shadow economy estimates) or, otherwise, included in our shadow economy estimate added to HFCE in step 2.
4. We assume that the passive shadow economy in a given country is related to the value of cash consumer payments on the country's territory. Yet, according to national accounts guidelines, HFCE for a given country includes (i) tourism expenditure of domestic consumers abroad and excludes (ii) tourism expenditure of foreign consumers on the territory of the analysed country. Therefore, we subtract (i) from the value obtained in step 3 and add (ii) to the resulting value, using the best available estimates of (i) and (ii) from the World Tourism Organisation included in the World Bank's database.
5. The value obtained in step 4 approximates the total value of monetary consumer payments on the territory of a given country (including the shadow economy). To estimate its cash component, we subtract from this value the estimates of the value of different types of non-cash consumer payments on the territory of the country (card payments, non-card electronic payments, payments with cheques). Data sources in this step depend on the analysed country, but they mainly include data from central banks, Mastercard and various market research reports. For all the analysed countries, data on the value of consumer card payments is available and included in this step. Yet, for some countries, data on the value of consumer non-card electronic payments and payments with cheques (which, however, are marginal in many countries) is not always available, in which case the respective value is not subtracted in this step. With respect to this,

[27] In other words, we assume that a x% decrease in consumer cash payments leads to both: (i) x% decrease in cash payments in the passive shadow economy and (ii) x% decrease in the registered cash payments. Please note that it may be a conservative assumption when considering the impact of a regulation targeted at a selected sector in which the role of the passive shadow economy is significantly higher than, on average, in a given economy.

[28] Most often, if any estimate of the shadow economy is available from a statistical office, it is presented as a percentage of the official GDP figure, without its breakdown into different components of GDP from the expenditure side, such as HFCE, investment, government expenditure and net exports.

two issues should be noted. First, in many countries, most of the value of non-card electronic payments is not related to consumer transactions but rather to business-to-business (B2B) payments or person-to-person (P2P) transfers, which do not (directly) influence processes that are investigated in this study. Second, available data show that in most countries the value of consumer card payments is significantly higher than the value of consumer non-card electronic payments.

Appendix 3. Estimation of the impact of the proposed measures on the value of card and cash payments

In this appendix, we present our methodological approach to the estimation of the impact of the selected measures on the value of consumer electronic and cash payments (as estimated in Appendix 2). When a considered solution positively affects the value of card payments, we assume that the resulting reduction in the value of consumer cash transactions leads to a proportional decline in the passive shadow economy.

Once we have calculated the impact of a given measure on the passive shadow economy, we can estimate the resulting change in government revenues (for methodological details, see Appendix 4). In addition, for some regulations we present the cost-benefit analysis that accounts not only for additional revenues, but also for the costs that the government may incur due to the introduction of the regulation. Furthermore, we sometimes consider the effects of a given solution in the short and long term.

For some countries, we analyse the effects of the considered regulations under two scenarios: (i) the **base scenario** representing the current estimated level of the passive and committed shadow economy; and a hypothetical (ii) **reduced committed shadow economy scenario** in which the committed shadow economy in the analysed country is reduced by 1/3, and an additional assumption is made that 40% of this contraction has converted into the expansion of the passive shadow economy (for a discussion on the relation between the passive and committed shadow economy, see chapter 1.3 in the Reports). Therefore, while in the reduced committed shadow economy scenario the total size of the shadow economy is reduced, its passive component increases. This additional reduced committed shadow economy scenario can be regarded as triggered by a regulatory package addressing not only the passive shadow economy, but also including solutions aimed at limiting the committed component.

Since our approach to estimating the impact of the selected measures on the structure of consumer payments varies with the considered solutions, we describe it separately for each of the investigated measures.

Obligation to operate POS terminals/improvement in the electronic payments acceptance (ensuring the right of consumers to pay electronically)

To simulate the impact of the obligation to operate POS terminals (or some other action that would obtain the same improvement in the electronic payments acceptance) on the value of card payments, we first estimate the number of POS terminals in different sectors in the analysed countries. We start with data on the total number of POS terminals (most often obtained from the central banks). In the first step, we break down this number into different sectors/merchant categories using (i) Mastercard's information on the number of POS terminals by merchant category, and, if (i) is not available, (ii) Mastercard's data on the sectorial breakdown of the number of payment card transactions performed at POS terminals.

In the second step, we assume that the number of terminals *per capita* in individual sectors, after the introduction of the regulation in the analysed countries, would be as high as it is currently in Australia (a country with one of the highest number of POS terminals *per capita* in the world, high popularity of electronic payments and for which we have obtained Mastercard's data on the sectorial breakdown of the number of POS terminals). We compute the gaps between a given country and Australia on a sector by sector basis. There are two factors that differentiate the gaps in the number of POS terminals in distinct sectors: (i) the size of a given sector and (ii) the level of its "saturation" with POS terminals. We compute the gaps only for the passive shadow economy sectors.[29] We also account for the differences in the consumption structure against Australia, since we expect relatively more (less) terminals in a given sector if its share in household consumption is higher (lower) than in Australia.

[29] Passive shadow economy sectors are described in Chapter 4 of the Reports.

In the third step, we aggregate the estimated sectorial gaps in the number of POS terminals to obtain the total gap for a given country and correct it for the number of unregistered enterprises. This correction is related to the fact that in practice the considered regulation will not cover unregistered companies (and their share in the total number of companies in some of the analysed countries is relatively high). Since direct estimates of the number of unregistered companies are not available, we multiply the estimated total gap in the number of POS terminals by the correction factor. We calculate this factor with the use of the following estimates:

(a) **The share of the monetary economy in the total GDP** (100% minus the non-monetary shadow economy and imputed rents expressed as percent of the total GDP, see step 2 in Appendix 1). We treat it as a part of the economy that would be covered by the regulation if no unregistered companies operated in the monetary economy.

(b) **The share of the monetary economy "covered" by the existing POS terminals in the total GDP** (to obtain it we multiply (a) by the share of the current number of POS terminals in the sum of the current number of terminals and the estimated gap in this area). In a country without POS terminals no part of the monetary economy is "covered" by POS terminals. On the other hand, in a country where the estimated gap is equal to zero, i.e. that has as many POS terminals *per capita* as Australia, all the monetary economy is "covered" by the existing POS terminals.

(c) **The share of the committed shadow economy in the total GDP** (see step 4 in Appendix 1), which reflects a part of the monetary economy where unregistered companies are present.

The correction factor is obtained using the following formula: $1-\frac{c}{a-b}$. The denominator of the fraction ($a-b$) approximates the share of the monetary economy not covered by the existing POS terminals in the total GDP. If it is equal to the share of the committed shadow in the total GDP (c in the numerator of the fraction), the correction factor is equal to zero. This corresponds to a situation in which all the companies without POS terminals in the economy are unregistered, so the considered regulation would have no impact on the number of POS terminals (since unregistered companies would rather not comply with the regulation). If c equals zero, the correction factor is equal to 1, which means that effectively we apply no correction to the estimated gap in the number of POS terminals. This corresponds to a situation in which the committed shadow is not present. In such a case, according to our approach, there are no unregistered companies in the economy, so no correction is needed. In general, if other things are constant, the higher the share of the committed shadow economy in the total GDP (i.e. the more unregistered companies are present in the economy), the lower the estimated gap in the number of POS terminals (after the applied correction), which may be reduced by the considered regulation. Consequently, in the reduced committed shadow economy scenario, in which the number of unregistered companies is lower, a regulation-driven increase in the number of POS terminals is higher.

In the fourth step, we translate the impact of closing the estimated gap (after the correction described above) in the number of POS terminals in *per capita* terms (i.e. a change in the variable **TERMINALS_PER_CAPITA**, see **Table A3.1**) on the ratio of card transactions value to consumer expenditure (**TRANSACTIONS_CONS**, see **Table A3.1**) using the econometric estimates from the following equation:

$$\boldsymbol{TRANSACTIONS\_CONS}_{it} = \alpha_i + \theta_t + \beta_{i,t}\boldsymbol{TERMINALS\_PER\_CAPITA}_{it} + \boldsymbol{\gamma x}_{it} + \varepsilon_{it} \quad \text{(A3.1)}$$

in which i, t denote the country and time subscript (year), respectively, $\boldsymbol{x}$ contains the **CARDS_ACTIVE_PER_CAPITA**, **GDP_REAL_PPS_PER_CAPITA** variables (see Table A1.1), α_i denotes the country effects, θ_t – the time effects (coefficients on period-specific dummies), ε_{it} is the error term while $\beta_{i,t}$ and $\boldsymbol{\gamma}$ represent the coefficients. As in the case of the currency demand model from substep 3.1 in Appendix 1, the $\beta_{i,t}$ is allowed to vary across countries and time, which is achieved through the interaction term of **TERMINALS_PER_CAPITA** with **GDP_REAL_PPS_PER_CAPITA**. For more details on the variables used and data sources see Tables A1.1 and A3.1. Table A3.2 presents the estimation results of two different variants of the model described in the equation (A3.1). If a variable is omitted in a given column, the corresponding coefficient equals zero. In each case, we used GLS estimator accounting for heteroskedasticity. The sample includes 62 countries: Argentina, Armenia, Australia, Austria, Belgium, Bolivia, Brazil, Bulgaria ,Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Dominican Republic, Egypt, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, India, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Latvia, Lithuania, Luxembourg, Macedonia, Malaysia, Malta, Montenegro, Netherlands, New Zealand, Nigeria, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Romania, Russian Federation, Saudi Arabia, Singapore, Slovak Republic, Slovenia, Spain, Sri Lanka, Sweden, Thailand, Tunisia, United Arab Emirates, United Kingdom, Uruguay, Vietnam, Yemen.

Table A3.1. Additional variables used in the regulatory impact analysis

Name of the variable	Description of the variable	Source(s)
TRANSACTIONS_CONS	The % share of card payments (made domestically and abroad, at POS terminals and including e-commerce) in household final consumption expenditure; annual data	World Bank (Global Payment Systems Survey, World Development Indicators), European Central Bank, local central banks and statistical officies, own calculations
CARD_FEES	Total variable fees (% of transaction value) charged to merchants due to card payments acceptance, including interchange fees and assessment fees; annual data	Mastercard
TERMINALS_PER_CAPITA	The number of POS terminals *per capita*; annual data	World Bank (Global Payment Systems Survey), local central banks, own calculations

Source: EY elaboration.

In Model 1 (column 1 in Table A3.2), there are no interaction terms, which should be incorporated into our model to account for diversity of countries included in the sample. Consequently, in column (2) we present a model containing an interaction of **TERMINALS_PER_CAPITA** and **GDP_REAL_PER_CAPITA** variables. The estimation results imply that the value of coefficient $\beta_{i,t}$ is larger, and hence the marginal benefits from expanding the POS terminal network are higher, in less developed countries and decreases as the GDP *per capita* grows over time. We treat the model in column (2) as the benchmark model used in our simulations.

Table A3.2. Econometric estimates of the determinants of the value of card payments in relation to consumption expenditure

	(1) Model 1	*(2)* Model 2
TERMINALS_PER_CAPITA	23.63*** (0.000)	**235.8***** **(0.000)**
GDP_REAL_PER_CAPITA ×TERMINALS_PER_CAPITA		**-4.705***** **(0.000)**
GDP_REAL_PER_CAPITA	0.533*** (0.000)	**0.553***** **(0.000)**
CARDS_ACTIVE_PER_CAPITA	3.746*** (0.000)	**4.233***** **(0.000)**
The number of observations	650	**650**
The number of countries	62	**62**

Notes: p-values in parentheses, * p<0.1, ** p<0.05, *** p<0.01. Individual effects for 62 countries (α_i) and 16 time periods (θ_t, years 2001-2016) are omitted for brevity of presentation, but included in each specification. All the models are estimated using the GLS estimator accounting for heteroskedasticity. The baseline model is the model in column (2).
Source: EY elaboration.

We assume that the value of cash payments is reduced by the amount equal to an increase in the value of POS transactions resulting from the introduction of the analysed regulation.

We use the modelling framework to compare short- and long-term effects of this regulation. Country-specific assumptions are described in respective country Reports.

Tax incentives for consumers

Our analysis of consumer reactions to monetary benefits related to card transactions is based on the microeconometric study of Arango et al.[30]. The authors analysed Canadian consumers' choice between different means of payment: cash, debit card and credit card. The proposed multinomial logit model makes it possible to calculate the substitution effects between different forms of payment as a result of the introduction of a 0.78% cash-back for credit card transactions (see Table A3.3). The model controls for both characteristics of each form of payment and additional characteristics of consumers (such as income, age, gender or opinions on particular means of payment).

Table A3.3. The impact of a 0.78% cash-back for credit card transactions (by transaction value, in CAD) on the changes in the probability of using various means of payment by consumers (in percentage points)

Transaction value (CAD)	Cash	Debit card	Credit card
5	-0.37 p.p.	-0.21 p.p.	0.58 p.p.
25	-1.82 p.p.	-1.78 p.p.	3.61 p.p.
50	-2.76 p.p.	-5.12 p.p.	7.87 p.p.
100	-2.1 p.p.	-10.71 p.p.	12.81 p.p.

Source: Arango C.A., Huynh K.P., Leonard Sabetti L. (2011), "How Do You Pay? The Role of Incentives at the Point-of-Sale", ECB Working Paper No. 1386., Table 12, p. 36. This table is consistent with Table 8 in Arango C., Huynh K.P., Sabetti L., "Consumer payment choice: Merchant card acceptance versus pricing incentives", *Journal of Banking & Finance*, vol. 55, 2015 (p. 139, "No selection" variant), but the results in the working paper are reported more precisely.

The estimated effect of introducing a 0.78% cash-back for credit card payments varies with the transaction value. However, we cannot differentiate between the effects of such cash-back for different transaction values in other countries, because we do not have information about the distribution of cash payments by transaction value in countries other than Canada. Nevertheless, we can calculate the cash-back effect for the average transaction value in Canada (which equals CAD 33), compute the resulting decrease in the probability of using cash and, finally, use this result to evaluate the effects for average transaction value in other countries.

Consequently, in the first step, we use a quadratic interpolation in order to calculate the reduction in the probability of using cash as a result of introducing a 0.78% cash-back for credit card payment.

[30] Arango C., Huynh K.P., Sabetti L., "Consumer payment choice: Merchant card acceptance versus pricing incentives", *Journal of Banking & Finance*, vol. 55, 2015, pages 130-141.

We have to use interpolation between transactions of CAD 25 and CAD 50 (the effects for the average transaction value of CAD 33 are not published). The interpolated effect, i.e. the reduction in the probability of using cash, for an average transaction value of CAD 33 equals 2.20 p.p.

In the second step, we use the following formula for computing the substitution effects:

$$\frac{\partial p_{ij}}{\partial x_i} = p_{ij}[\beta_j - \sum_{l=1}^{m} p_{il}\,\beta_l], \text{ (A3.2)}$$

where p_{ij} stands for the probability of consumer *i* to choose the means of payment *j* (credit card - CC, debit card – DC, or cash – CASH); x_i is the observed factor that we change (in this case, a cash-back reward in CAD); β_j is the estimated parameter from the logit model measuring the impact of cash-back on the perceived attractiveness of the means of payment *j*. Note that l indexes all possible methods of payment, while *m* stands for the total number of the considered means of payment.

After some basic transformations and using the assumption of the authors of the study that $\beta_{CASH} = \beta_{DC} = 0$ (because there is no cash-back for cash and debit card transactions in the Canadian sample), the substitution effect for cash payments simplifies to:

$$\frac{\partial \mathrm{p}_{iCASH}}{\partial \mathrm{x}_i} = -\mathrm{p}_{iCASH}\mathrm{p}_{iCC}\beta_{CC}. \text{ (A3.3)}$$

Note that while in Arango et al.[31] debit card payments are not rewarded with any cash-back, in our study we consider regulations that affect both debit and credit card transactions. In order to analyse the effect of a cash-back introduced for both kinds of card payments, we consider an increased probability of using credit cards equal to $p_{iCC}^{*} = p_{iCC} + p_{iDC}$ and subsequently reduce the probability of using debit cards to zero (so that the sum of probabilities of using different means of payment still equals 1). Such a transformation allows us to treat the category "credit cards" as if it included both credit and debit cards. Consequently, we are now able to investigate the effect of a cash-back for all sorts of card payments. We achieve this by substituting p_{iCC}^{*} for p_{iCC} in equation (A5.6), which allows us to obtain the strengthened substitution effect for cash:

$$\begin{aligned}\frac{\partial p_{iCASH}}{\partial x_i}^{*} &= -p_{iCASH}p_{iCC}^{*}\beta_{CC} \\ &= -p_{iCASH}(p_{iCC} + p_{iDC})\beta_{CC} \\ &= -\frac{(p_{iCC} + p_{iDC})}{p_{iCC}}p_{iCASH}p_{iCC}\beta_{CC} \\ &= \frac{(p_{iCC}+p_{iDC})}{p_{iCC}}\frac{\partial p_{iCASH}}{\partial x_i}. \text{ (A3.4)}\end{aligned}$$

In the study of Arango et al.[32], for the average transaction value of CAD 33, the probability of using cash declines by 2.20 p.p. as a result of a cash-back awarded exclusively to credit card transactions ($\frac{\partial p_{iCASH}}{\partial x_i}$). The shares of the debit card (p_{iDC}) and credit card (p_{iCC}) payments in consumer transactions equal 37% and 30%, respectively. With this information, we can use the formula (A3.4) to conclude that the introduction of a 0.78% cash-back for both debit card and credit card transactions would reduce the probability of using cash by: (30+37)/30*2.2 p.p.=4.90 p.p.[33] Taking into account that the probability of using cash equals 33%, a 4.90 p.p. decrease of this probability is equivalent to a reduction in the demand for cash by 4.9/33=14.9%. This effect should be rescaled accordingly if the considered incentive for card payments differed from a 0.78% cash-back analysed by Arango et al.[34] For example, if the incentive equalled 0.5% of the transaction value, it would lead to a 0.5/0.78*14.9%=9.5% reduction in the demand for cash. The same relative decrease in the demand for cash (i.e. 9.5%) would apply to each analysed country – for instance, if the probability of using

[31] Arango et al (2015), *op. cit*.

[32] Arango et al (2015), *op. cit*.

[33] We use formulas for an individual consumer *i* in order to make them consistent with the original study of Arango et al. (2015), *op. cit*.; yet we assume that the correcting factor $\frac{(p_{iCC}+p_{iDC})}{p_{iCC}}$ is the same for all consumers and all the transactions worth CAD 33.

[34] Arango et al (2015), *op. cit*.

cash amounts to 30% in a given country, it would drop after the implementation of the regulation by 9.5%*30%=2.9 p.p., while the value of card transactions would increase by 2.9/(100-30)=4.1%.[35]

The reduction in cash transactions value leads to a proportional decrease in the passive shadow economy. The respective increase in the registered economic activities results in higher government revenues. The passive shadow economy is larger in **the reduced committed shadow economy scenario** than in **the base scenario** and so are the government benefits stemming from the introduction of the considered solution.

On the other hand, the regulation also entails costs in the form of reduced government revenues per registered card transaction, due to deducting a fraction of the tax burden. The actual level of costs and benefits, and the resulting net effect for the government revenues, is highly country-specific (due to heterogeneity in tax rates, as well as levels and structures of consumer payments). Consequently, the optimum level of the consumer incentive varies with the country analysed. In addition, the optimum result also depends on whether we consider **the base scenario** or **the reduced committed shadow economy scenario.**

Tax incentives for merchants

We analyse the effect of changes in merchant's costs of accepting card payments through the econometric estimation of the equation:

$$\boldsymbol{TRANSACTIONS\_CONS}_{it} = \alpha_i + \beta \boldsymbol{CARD\_FEES}_{it} + \boldsymbol{\gamma x}_{it} + \varepsilon_{it} \quad \text{(A3.5)}$$

where i, t denote country and time subscript, respectively, $\boldsymbol{x}$ stands for other factors affecting the value of transactions, namely **CARDS_ACTIVE_PER_CAPITA** and **GDP_REAL_PPS_PER_CAPITA** (for details on the variables used and data sources see Tables A1.1 and A3.1) and a linear trend variable, while β and γ represent the corresponding coefficients. Such specification implies that the impact of a tax incentive on the share of card transactions in consumer expenditures is proportional to the tax incentive itself, i.e. a 2% tax relief implemented via fee reduction leads to twice as high an increase in **TRANSACTIONS_CONS** as in the case of a 1% tax relief. There are 29 countries in the sample: Australia, Austria, Bulgaria, Croatia, Czech Republic, Denmark, Egypt, Finland, France, Germany, Greece, Hungary, India, Indonesia, Italy, Japan, Norway, Pakistan, Poland, Portugal, Russian Federation, Saudi Arabia, Slovak Republic, Slovenia, Spain, Sweden, United Arab Emirates, United Kingdom, Vietnam. The sample covers the 2010-2016 period. The panel is unbalanced and suffers from relatively short time dimension (as compared to the number of countries); furthermore, the cross-country variability ("between") dominates over the variability in time ("within") for the key variables (especially **CARD_FEES)**. For this reason, we do not use fixed effects in either dimension (time or space), and focus on random individual effects and use linear trend instead. GLS estimation results (column (3) in Table A3.4) are chosen as the baseline model due to panel-specific heteroskedasticity.

Table A3.4. Econometric estimates of the impact of merchant costs on the value of card transactions

Estimator	(1) OLS	(2) Random Effects (RE)	(3) RE GLS
CARD_FEES	-0.0514** (0.020)	-0.0101 (0.260)	**-0.0271*** (0.000)**
GDP_REAL_PER_CAPITA	0.327*** (0.002)	0.555** (0.011)	**0.200*** (0.000)**
CARDS_ACTIVE_PER_CAPITA	15.46*** (0.000)	11.48** (0.022)	**18.48*** (0.000)**
Time	0.866	1.033***	**1.057*** **

[35] Note that 9.5% reduction in the demand for cash applies to every country in our analysis, while 30% in this example stands for a country-specific share of cash in consumer transactions, which varies among countries.

	(0.116)	(0.000)	**(0.000)**
The number of observations	129	129	**129**
The number of countries	29	29	**29**

Notes: p-values in parentheses, * p<0.1, ** p<0.05, *** p<0.01. The GLS estimator accounts for panel-specific heteroskedasticity. The baseline is the model in column (3).
Source: EY elaboration.

We assume that the value of cash payments is reduced by the same amount by which the value of POS transactions is increased as a result of the introduction of the analysed regulation. The difference between **the base scenario** and **the reduced committed shadow economy scenario** is related solely to the fact that the passive shadow economy is higher in the latter scenario. As in the case of tax incentives for consumers, an analysis of costs and benefits is conducted and the optimum rates of tax incentives are provided for each country separately.

Obligation to make an electronic payment of wages and salaries

We estimate the impact of this regulation on the value of non-cash payments in a few steps. First, we use the World Bank's survey data for the analysed countries (Global Findex Database) on the percentage of population that received their wages in 2017 in cash only[36] to compute the count of such people. Second, we assume that all the persons employed in the shadow economy (unregistered employees) receive their wages in cash. Third, using the estimates of the number of unregistered employees[37], we calculate the number of people who are officially employed and receive their wages in cash. This is the target group of the regulation. Next, we make four additional assumptions:

(1) the target group earns the average wage;

(2) the target group currently spends all the received wages in cash[38];

(3) the savings rate of the target group is equal to the savings rate of all the households in a given country (if such data is available) or to 10% (otherwise);

(4) if the target group received wages into a bank account or pre-paid card, they would spend the same share of this money in the form of non-cash payments as the total population of the current payment card or account users in a given country. This share is computed as the ratio of domestic card payments value to the sum of: (a) the value of card payments (both domestic and done abroad) and (b) the value of cash withdrawals at ATMs (both domestic and done abroad).[39]

[36] This percentage should be treated as a conservative estimate for our purposes. This is because our calculations account for the impact of the regulation on the employees that receive wages in cash only, while there are also some employees who receive just part of their wages in cash. Accounting for the impact of the regulation on the latter would increase the calculated effects, respectively (however, there is no sufficient data available to do so).

[37] Note the difference between unregistered employees discussed here and the unregistered employment (including also the self-employed) estimated in step 4 in Appendix 1.

[38] Probably there are some people who earn their wages in cash and deposit this money into a bank account. Nevertheless, we may expect that it is not common.

[39] The calculation of this share has been conducted for the Mastercard's payment card users with the use of data provided by Mastercard. We assume it is representative for all payment card users in a given country. For some countries, the estimates of the value of other kinds of consumer cashless payments (non-card) are available. If the sum of value shares of card and non-card cashless payments in consumer payments was X% higher than the value share of card payments in consumer payments, then we made the estimated share used in point (4) also X% higher.

We obtain the value of additional cashless payments generated by the regulation for each country by multiplying: the size of the target group, the average wage, household consumption rate (one minus savings rate) and the share of consumer spending in the form of non-cash payments (as discussed in point (4)). This value also indicates the value of consumer cash transactions being crowded out by new cashless payments.

In addition, in the **reduced committed shadow economy scenario**, we assume that the number of unregistered employees would decline by the same percentage as the committed shadow economy and these newly registered employees would be subject to the considered regulation.

It should also be noted that our estimates show the effect of the scenario in which there is a full compliance with the regulation (it does not take into consideration that some entities may violate the law and/or that there may be some exemptions from the regulation, etc.).

Obligation to make an electronic payment of social security benefits

The estimation of the impact of this regulation on the value of non-cash payments is similar to the approach applied to the obligation to make an electronic payment of wages and salaries (see above).

First, we use the data on the total value of social security benefits/social expenditure in a given country. When such data is available, we focus our analysis only on the part of the social transfers that is obtained by individuals in the monetary form (so we omit, e.g., various types of in kind benefits); otherwise we consider the entire social spending category. For that purpose, we use various national sources and international databases (e.g. OECD database). Second, we use the World Bank's survey data (Global Findex Database) on the percentage of government payments (transfers) recipients in the analysed countries who received these transfers in cash only[40] in 2017. We also take into account the household savings rates and the share of consumer spending in the form of non-cash payments, as discussed in point (4) of the description of the obligation to make an electronic payment of wages and salaries. For each country we multiply: the total value of the considered social security benefits, the share of cash recipients in the total number of recipients of government transfers, the households consumption rate (one minus savings rate) and the share of consumer spending in the form of cashless payments, which allows us to obtain the value of additional cashless payments generated by the introduction of the regulation.

Threshold for consumer cash payments

In order to estimate the effect of introducing an upper limit for consumer cash payments on the passive shadow economy, we use the data on the distribution of consumer cash payments in Poland, based on the consumer payment diaries and provided by the courtesy of the Polish central bank (NBP, see **Table A3.5**).

In a vast majority of the analysed countries, no studies have been conducted that could be helpful in determining the structure of consumer cash payments by the size of transactions. In some of them, e.g. Australia, such studies were carried out but the data remains confidential. Therefore, our approach is based on the NBP data, while certain assumptions are adjusted to the country-specific macroeconomic situation in other countries.

Table A3.5. Distribution of consumer cash payments in Poland according to payment diaries

	Value of the threshold in PLN	Value of the threshold in USD	Value of cash payments below the threshold (%)	Value of cash payments above the threshold (%)	Relative decrease in the passive shadow economy (%)
1	20	5.1	15	85	83.9
2	50	12.7	46	54	50.5

[40] This percentage should be treated as a conservative estimate for our analysis. This is because, due to limited data availability, our calculations account for the impact of the regulation on the people that receive government transfers in cash only, while there are also some people who receive just part of transfers in cash (or one type of transfers in cash and other electronically). Accounting for the impact of the regulation on the latter would increase the calculated effects, respectively.

3	100	25.4	70	30	24.7
4	150	38.0	81	19	12.9
5	200	50.7	85	15	8.6
6	300	76.1	93	7	0.0

Note: Applied USD/PLN exchange rate: 3.94.
Source: NBP data

Data on the distribution of consumer cash payments shows that the value of consumer cash payments is marginal for higher transaction values. Moreover, above a (relatively) high threshold of the transaction value, there should be almost no passive shadow economy, because one can expect that consumers tend to demand receipts for more expensive goods or services in order to obtain a warranty. Obviously, there are high-value cash payments in the committed shadow economy. However, these would remain unaffected by the considered regulation, as both parties benefiting from this kind of activity would continue to use cash in order to avoid reporting the transaction. Consequently, establishing a threshold at a high level would largely lead to replacing the already reported cash payments with card payments and thus would have little, if any, impact on the shadow economy.

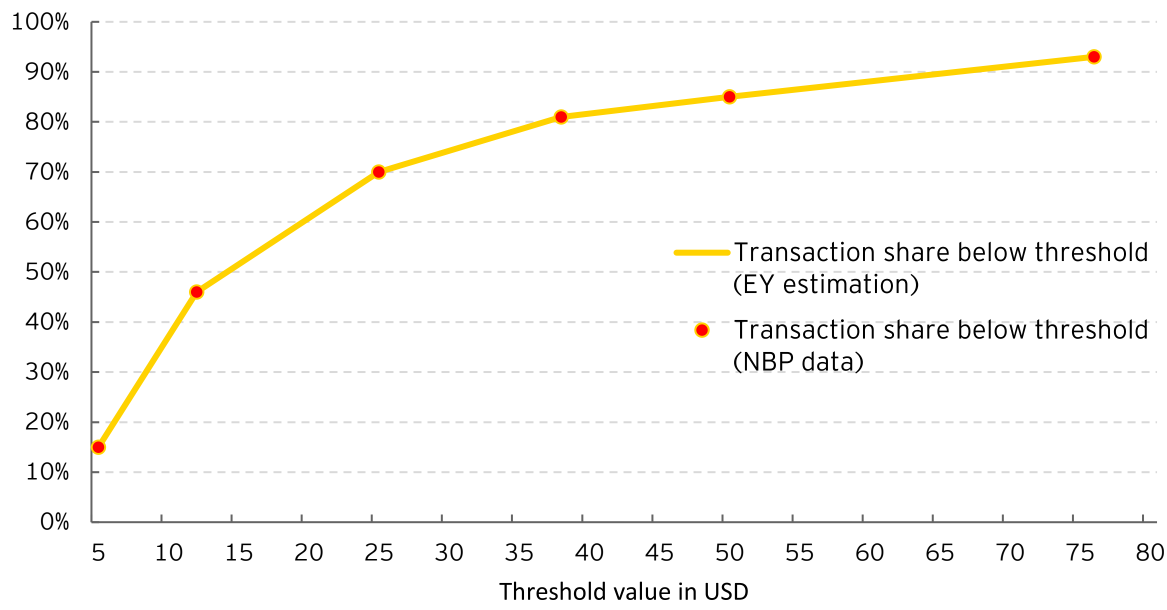
We account for this fact by assuming that above the threshold no. 6 (the highest threshold value considered in the original data of the Polish central bank, see **Table A3.5**), there are no passive shadow economy transactions. Such an assumption makes our impact assessment more conservative. It implies that in each country the top 7% of consumer cash transactions (in terms of their value) remain unaffected by the regulation. In other words, the relative fall in cash usage, corresponding to the value of cash payments above a given threshold (see column 4 in **Table A3.5**), is higher than the relative decrease in the passive shadow economy (column 5). This relative impact of each threshold on the passive shadow economy is calculated according to the following formula:

$$Percentage\ decrease\ in\ the\ passive\ shadow\ economy = \frac{(\%\ of\ value\ of\ cash\ payments\ above\ the\ threshold - 7\%)}{(100\% - 7\%)} \times 100\% \quad (A3.6)$$

We assume that this percentage change is the same for all the analysed countries for certain levels of the respective thresholds. To calculate the thresholds for other countries, we adjust the Polish values (expressed in USD) by differences in nominal household final consumption expenditure *per capita* between the analysed country in 2016 and Poland in 2012 (a year in which the NBP study was conducted).[41] In other words, we assume that the distribution of cash payments is analogous in all the analysed countries, after adjusting for the differences in their household consumption levels, which approximate the differences in the value of their consumer cash payments.

Chart A3.1. Share of consumer cash transactions (in %) below various thresholds – NBP data and EY estimation

[41] In the case of countries where data on the household final consumption expenditure in 2016 is missing, we assume that its share in GDP in 2016 is the same as in the latest period for which the data is available, i.e. 2012 for Jordan and 2015 for Myanmar.



Source: EY calculations, NBP data

We calculate thresholds expressed in national currencies (round values). The round thresholds in PLN (the currency of the NBP study) do not translate into the round values of thresholds in other currencies. Consequently, we cannot directly use the cumulative shares of cash transactions below a given value, as specified in **Table A3.5**. We therefore have to estimate the respective cumulative shares of transactions, associated with the rounded thresholds in local currencies. We achieve it through the linear interpolation between the dots presented in **Chart A3.1** based on the NBP data.

We assume that the structure of the passive shadow economy transactions by their size is the same in **the base scenario** and **the reduced committed shadow economy scenario**. Therefore, additional passive shadow economy transactions that are generated as a result of the committed shadow economy reduction are distributed proportionally to consumer cash payments under **the base scenario**. Consequently, the imposition of a given threshold results in the same relative decrease in the passive shadow economy size in both scenarios.

Promotion of electronic payment devices that enable contactless payments

We assess the impact of the promotion of electronic payment devices that enable contactless payments on the value of electronic (card) and cash payments using a simulation approach. We consider the effects of the scenario in which all contact payment cards and POS terminals in the analysed country would be replaced with ones that enable contactless payments. We concentrate on payment cards and POS terminals, since such electronic payment devices are the most popular in many countries around the world and availability of the data on their numbers (required for the econometric analysis) was the highest for us. However, replacing contact cards and POS terminals (or alternative devices currently used for contact payments) with other devices that enable contactless payments (e.g. in some countries mobile phones may be already used for contactless payments and their acceptance) should generate similar effects per replaced device.

We conduct our calculations in the following steps. First, for the analysed country and year, we calculate the product of the shares of: (1) contactless cards in the total number of payments cards and (2) contactless POS terminals in the total number of POS terminals (**PRODUCT_CONTACTLESS_SHARES**). We use Mastercard's data (then it covers only Mastercard branded cards and POS terminals that accept payments with such cards) or other available sources (e.g. the central bank, association of banks, etc.). We focus on the product of the shares (instead of analysing the two shares separately), reasoning that contactless card payments require both the contactless cards and contactless POS terminals. For example, if all POS terminals were contactless but no cards had this function, the calculated product would be equal to 0 (the lowest possible value) and no contactless card payments would happen. In addition, by focusing on the product of the discussed shares, instead of having two separate variables, we later avoid potential issues related to their collinearity in our econometric analysis.

Second, we calculate the gap between the product from step 1 and its maximum possible value of 1 (the value of 1 reflects the situation in which all payment cards and terminals would become contactless).

Third, we compute how the elimination of the gap calculated in step 2 would change the share of the value of card payments in household final consumption expenditure (**TRANSACTIONS_CONS**, for details see table A3.1). For this purpose, we use the econometric estimates from the following equation:

$$\boldsymbol{ln}(\boldsymbol{TRANSACTIONS\_CONS})_{it} = \alpha_i + \beta \cdot \boldsymbol{PRODUCT\_CONTACTLESS\_SHARES}_{it} + \boldsymbol{\gamma} \cdot \boldsymbol{x}_{it} + \varepsilon_{it} \quad \text{(A3.7)}$$

in which i, t denote the country and time subscript (year), respectively, $\boldsymbol{x}$ contains the **CARDS_ACTIVE_PER_CAPITA**, **TERMINALS_PER_CAPITA** and **GDP_REAL_PPS_PER_CAPITA** variables (see Table A1.1 and A3.1 for more details) as well as **UK_DUMMY** (variable which was equal to 1 for the United Kingdom from 2013 and was equal to zero for all other periods and countries to eliminate an outlier in the distribution of the error term), α_i denotes the country effects, ε_{it} is the error term, while β and $\boldsymbol{\gamma}$ represent the coefficients.[42] The sample includes years 2000-2018 (for most countries 2000-2017) and the following countries: Albania, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Finland, France, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Malta, the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and United Kingdom. For these countries time series data on our variable of interest (**PRODUCT_CONTACTLESS_SHARES**) was shared with us by Mastercard.

We use the natural logarithm (***ln***) of the dependent variable, so an absolute change in **PRODUCT_CONTACTLESS_SHARES is related with** a *percentage* change of **TRANSACTIONS_CONS**. Such relation seems reasonable, since in the applied setup the resulting *absolute* change of **TRANSACTIONS_CONS is high (low) in countries with the currently high (low) value of card payments, which is to large extent an effect of the currently high (low) number of payment cards and POS terminals (contact and contactless ones) in the country.** Therefore, in most cases, the higher the current overall number of payment cards and POS terminals in a country, the higher absolute increase in the value of card payments would be obtained by replacing a given share of cards and/or POS terminals with ones that enable contactless payments. Using the described approach we separate the effect of "contactlessness" of payment cards and POS terminals from the effects of the increase in their overall numbers (which were controlled in our econometric model).

Table A3.6 presents the estimation results of two different variants of the model described in the equation A3.7. We choose fixed effects panel estimators, since there was a lot of variability of the independent variable of our interest (**PRODUCT_CONTACTLESS_SHARES**) in time (in years 2000-2017 many countries moved from the state with almost no contactless cards/POS terminals to very high shares of them in their total numbers) and because the result of Hausman test[43] indicate such approach. For fixed effects models we consider the general least squares and panel corrected standard errors estimators which are often used for the datasets that are similar to ours. We choose the former as our baseline model, since according to the research of Moundigbaye et al.[44], it should perform better in terms of efficiency for our data.

[42] We do not include **CARD_FEES** variable in the model, since this variable was available to us for different countries and less time periods than the variable of our interest **(PRODUCT_CONTACTLESS_SHARES),** so its inclusion would significantly reduce the number of observations used for our econometric estimation.

[43] Hausman J.A. (1978), "Specification tests in econometrics," Econometrica 46, pages 1251–1271.

[44] Moundigbaye M., Rea W.S, Reed W.R (2017), "Which Panel Data Estimator Should I Use?: A Corrigendum and Extension," Working Papers in Economics 17/10, University of Canterbury, Department of Economics and Finance.

Table A3.6. Econometric estimates of the impact of contactless payment cards and POS terminals on the value of card transactions

Estimator	(1) **Fixed effects general least squares**	(2) Fixed effects with panel corrected standard errors
PRODUCT_CONTACTLESS_SHARES	**0.347*** (0.000)**	0.398* (0.068)
CARDS_ACTIVE_PER_CAPITA	**0.683*** (0.000)**	0.920*** (0.000)
TERMINALS_PER_CAPITA	**17.084*** (0.000)**	16.969*** (0.000)
GDP_REAL_PER_CAPITA	**53.607*** (0.000)**	51.642*** (0.000)
The number of observations	**248**	248
The number of countries	**26**	26

Notes: p-values in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$. Both estimators account for heteroscedasticity and AR(1) autocorrelation with the coefficient of AR(1) process common to all the panels. UK-DUMMY and fixed effects were omitted in the table. The baseline is the model in column (1).
Source: EY elaboration.

In the last step, we translate the percentage growth in the value of card payments into their absolute growth which, in turn, is equal to the value of the crowded out cash payments. After that we calculate the resulting decrease in the size of the passive shadow economy as well as the related increase in government revenues. These effects should be deemed conservative, since we do not account for the fact that contactless payments relatively often crowd out cash payments of lower value among which the passive shadow economy transactions are more prevalent. Country-specific assumptions are described in respective country Reports.

Appendix 4. Impact of reducing the passive shadow economy on government revenues

In this appendix, we describe how we estimate the impact of the contraction of the passive shadow economy (and the resulting increase of registered transactions value) on the growth of the tax base that translates into additional government revenues. The results of these calculations are used in the impact assessment of the regulatory measures considered in Appendix 3 and the Reports.

In our analysis, we focus on two major categories of tax revenues related to companies' reported activity - Value Added Tax (VAT), or a similar tax category for some countries, and Corporate Income Tax (CIT).

Value Added Tax[45]

To estimate the value of additional VAT revenues due to a reduction of the passive shadow economy, we multiply the value of newly registered consumer expenditure by an estimated theoretical VAT rate. Most countries have different VAT rates for certain categories of goods and services, which means that we need to take into account the sectorial structure of the shadow economy to assess this rate in every country. We assume that the structure of consumer cash expenditure in the passive shadow economy is the same as the representative structure of household consumption, as reflected by the weights from the basket of consumer goods and services used in calculation of the CPI inflation (based on Household Budget Surveys). Accordingly, we calculate the theoretical VAT rate for the passive shadow economy transactions as a weighted average of official (standard or reduced) VAT rates applied to different goods/services in the economy, computed as if all the transactions were reported. We also take into account that some services or goods might be exempted from the VAT.

We use the weights derived from the household consumption expenditure data on the lowest available level of aggregation and assign the relevant official VAT rate to each category. However, in some cases goods/services in a given sector are subject to more than one VAT rate (e.g. in many countries there are VAT exemptions for very specific types of basic goods, such as bread, rice etc.). In such situations, it is not feasible to compute the detailed structure without additional data and we use either the lower rate or the mean value of the rates (the latter only if there is evidence that the higher rate prevails). This is consistent with the conservative approach adopted in this study. Consequently, it is more likely that we underestimate rather than overestimate a potential increase in VAT revenues as a result of the analysed measures.

Unless the regulation focuses on specific sectors, we assume that each sector is influenced by the analysed measure proportionately to its share in household consumption expenditure. This implies that the structure of activities that become registered due to the introduced regulatory measure is the same as the structure of household consumption expenditure.

Corporate Income Tax

If a decrease in the passive shadow economy influenced registered revenues, but not the costs incurred by businesses (which were already reported prior to a regulatory change under consideration), a growth in reported transactions would be equal to an increase in corporate profits plus VAT that is included in the price of these transactions.

[45] In this section we refer to consumption taxes that can have the form of Value Added Tax (VAT) or the General Sales Tax (GST). The general idea of tax income lost due to the existence of the shadow economy remains the same for either of these taxes.

However, to be on the conservative side, we treat the value of the transactions shifted from the shadow economy to the registered activity (without VAT) as an increase in their operating surplus[46] rather than profits of companies. It should be noted that the operating surplus is a broader category than the tax base for CIT. Therefore, we allow for the possibility that businesses had not reported all the costs before the new transactions were registered and that these costs can now be subtracted from the operating surplus. This, in turn, leads to a smaller increase in the tax base for CIT than the value of the newly registered transactions (less VAT).

In our approach, we estimate the "effective" CIT rate on the basis of data on revenues from taxes on corporate profits, usually published by the Ministry of Finance. We begin with an evaluation of the operating surplus of companies. To this aim, as a starting point, we use the total value added in the economy and subtract from it the value of registered employees' compensations, the imputed rents (not subject to CIT) (see Appendix 1, step 2) and the value of the shadow economy included in the official estimates of the GDP (if such data is not available, we use our estimates of the shadow economy). Next, we divide government CIT revenues by the obtained value of the operating surplus to calculate the "effective" CIT rate. In the last step, we calculate the product of the newly registered transactions (less VAT) and the effective CIT rate. This is our proxy for additional CIT revenues associated with a contraction of the passive shadow economy.

Lost government revenues due to existence of the passive shadow economy

In the Reports, we present (1) lost government revenues due to existence of the passive shadow economy. We calculate them by estimating the impact of reducing the passive shadow economy to zero for VAT (or a similar tax category for some countries) and CIT. In some countries, we also estimate (2) lost government revenues due to the existence of the committed shadow economy. In such a case, we use the same approach and the same VAT and CIT rates for the committed shadow economy as described above for the passive shadow economy. The sum of (1) and (2) constitutes lost government revenues due to the existence of the cash shadow economy, which we also present in the Reports for some countries.

[46] In the national accounts methodology used in the calculation of the GDP, there is a category of gross operating surplus, which can be viewed as close to the aggregated value of individual companies' EBITDA. However, the data on the value of gross operating surplus is available only in some of the analysed countries and, as a result, in our approach we use the notion of "operating surplus", calculated as the total gross value added less registered employees' compensations, imputed rents and the shadow economy included in the statistical office estimates of GDP.

EY | Assurance | Tax | Transactions | Advisory

About EY

EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services we deliver help build trust and confidence in the capital markets and in economies the world over. We develop outstanding leaders who team to deliver on our promises to all of our stakeholders. In so doing, we play a critical role in building a better working world for our people, for our clients and for our communities.

EY refers to the global organization and/or one or more of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients.

EY analysis was conducted based on publicly available data, data provided by the Client as well as EY insights. EY has not performed any audit/assurance of the data used in the analysis.

EY analysis is based on statistical and/or econometric models. For this reason, the results of analyses should be taken to be an approximation to any true relationships, which is conditional on the method of analysis, the data used and any expert assumptions. If any of the assumptions, the data used or the methods of analyses are revised, the results of the project analyses may change relative to the results set out in this document. EY assures the highest degree of professional care and conduct but regard must be taken of the uncertainties inherent in economic modeling.

Neither EY nor the employees of EY can be held responsible for the activities taken, or the lack of such activities, based on the information included in this document. EY accepts no loss arising from any action taken or not taken by anyone using this publication.

EY responsibility for the contents of this document is solely to the Client. EY takes no responsibility whatsoever to third parties using this document. None of the services or any part of the document constitute any legal opinion or advice.

For more information about our organization, please visit ey.com.

EY, Rondo ONZ 1, 00-124 Warszawa

This material has been prepared for general informational purposes only and is not intended to be relied upon as accounting, tax, or other professional advice. Please refer to your advisors for specific advice