

The distributional impact of reforms to direct and indirect tax in Mexico

Methodological Issues and Approach

Laura Abramovsky[♦], Orazio Attanasio^{*}, Carl Emmerson[♦],
David Phillips[♦], and Hector J. Villarreal Paez[^]

November 2010

Abstract

In 2010, the Mexican government implemented a fiscal tightening through an increase in VAT, the financial deposit tax, and a temporary increase to the top rate of income tax. This differed significantly from earlier proposals for a more significant tightening and reform that were rejected by the Congress. When assessing both the proposed and approved reforms, an important element of any appraisal is to ascertain the distributional impact of the tax changes. This paper lays out the methodological issues involved in such an exercise and describes the approach the authors will take in their analysis. It is the first paper of a study aimed at building capacity for the distributional analysis of tax reform in Mexico through, in part, the development of a micro-simulation tool, MEXTAX.

The Mexican context presents a number of challenges for such distributional analysis. First, is the high degree of economic informality and tax evasion meaning that the actual tax burden is significantly less than the legislated burden. Second is that the quality of income data in available surveys are poor relative to developed economies (although superior to that available in most developing countries). Finally, there is a relative paucity of estimates of the behavioral response to taxation in Mexico.

The methodology that shall be adopted for this study will make heavy use of robustness analysis to test the sensitivity of results to different assumptions about informality, tax evasion, missing income and behavioral response. The response of consumers to changes in VAT will be assessed using a consumer demand model, and an attempt will be made to analyse the decision of whether to work formally or informally. Rather than attempt to simultaneously model all behavioral responses simultaneously, different margins will be modelled or simulated independently.

The need for future work on behavioral response and suggested fruitful avenues for research are highlighted.

[♦]Institute for Fiscal Studies. Email addresses: labramovsky@ifs.org.uk, david_p@ifs.org.uk (corresponding author) and cemmerson@ifs.org.uk.

^{*} University College London, Institute for Fiscal Studies and NBER (o.attanasio@ucl.ac.uk).

[^] Centro de Investigacion Economica y Presupuestaria (hectorj.villarreal@gmail.com)

Contents

1. Introduction	3
2. Issues in tax microsimulation	6
2.1 Incidence assumptions	6
2.2 Behavioral response	7
2.3 Dynamic considerations	8
2.4 Informality	8
2.5 Data quality and availability	9
2.6 IFS analysis of UK reforms	11
3. Tax microsimulation in Mexico	13
3.1 Previous efforts	13
3.2 Our proposals, in general	15
3.3 The tax microsimulator	17
3.4 Behavioral modelling	23
3.5. Robustness analysis	27
3.6 Structure of the papers	31
3.7 Collaboration and capacity-building	32
4. Avenues for future research	32
5. Conclusion	35
Bibliography	37
Appendices	
A. A description of the Mexican tax system	40
B. The planned and implemented tax reforms	50
C. The MEXTAX data creation programs	53
D. Descriptive statistics	61

1. Introduction

In 2009, in Mexico, the government debt to gross domestic product (GDP) ratio stood at 35.6%, while the government deficit accounted for 2.32% of GDP. While these figures are low relative to the position of several developed countries, they hide a substantial imbalance: government revenues from taxation account for only 9.5% of GDP, while expenditure stands at 26.1% of GDP.¹ The difference between these figures is mainly covered by oil revenues, which therefore play an important role in guaranteeing the long-term solvency of the Mexican government. However, given the volatility of oil prices and the fact that proven reserves of Mexican oil are expected to last less than 10 years, there is an urgent need to consolidate government finances, both in terms of expenditure and in terms of revenue.

In terms of revenue, one of the main reasons for the low level of taxes to GDP is the large role of the informal economy, which evades taxation almost completely. Another reason is the existence of loopholes in the Mexican tax system that allows big businesses and the self-employed to legally avoid paying taxes. On the expenditure side, one concern is whether the way some programmes are designed constitutes an incentive to informality (see Levy (2008)).

In addition to these structural features of the Mexican economy, in 2009 Mexico's GDP shrunk by about 6.5%, which led to a decrease in non-oil tax revenues of 11.5%. Oil revenues also experienced a decrease of 21.4% due to a reduction of international oil prices and in the production of oil.²

In 2009, in response to the short-run reduction in fiscal revenues, and the perceived need to consolidate the fiscal budget in the long-run, the Mexican government implemented a modest fiscal tightening through an increase in the rate of VAT of 1%³, an increase in some duties, an increase in the financial deposit tax from 2% to 3%, and a temporary increase in the top rate of income tax from 28% to 30%. The Mexican Congress rejected more radical proposals for larger increases in duty rates, the introduction of a comprehensive 2% VAT on all goods (including those currently not covered), and increases in regulated prices. When assessing fiscal reforms such as these, an important element of the appraisal is to ascertain the distributional impact of the reforms. This paper lays out the methodological issues involved in such an exercise and describes the approach the authors will take in their analysis. It is the first paper of a study that aims to build capacity amongst researchers and practitioners for the analysis of the distributional and behavioral consequences of tax reform in Mexico (and eventually elsewhere in Latin America). Ultimately it will involve the creation of an analytical tool - MEXTAX - that will be able to be used to assess the distributional impact of past, future and hypothetical reforms of income tax and indirect taxes.

Quantitative analysis of the distributional impact of tax and benefit changes is common in many developed countries, and the results of such analyses are important in informing both the policy-making process and the public debate about the merits of such policies. For instance, in the United Kingdom (UK), the Government makes use of tools called the Intra Government Tax-Benefit Model (IGOTM), and the Policy Simulation Model (PSM) to analyse the distributional

¹ Bank of Mexico, Annual Report 2009 (see Cuadros 6, 7 and 18).

² Bank of Mexico, Annual Report 2009 (see Cuadro 7).

³ The main VAT rate increased from 15% to 16%, and the rate at which transactions subject to VAT are taxed in areas bordering the United States increased from 10% to 11%.

impact of changes to tax and income-related welfare programs.⁴ The independent Institute for Fiscal Studies (IFS) maintains a detailed simulator called TAXBEN which is used to produce timely analysis of proposed tax and welfare changes (see, for instance, Browne and Phillips (2010) and Adam et al (2010)), and in academic research (see Meghir and Phillips (2010)). Tax and welfare simulators have also been developed for continental European nations (EUROMOD)⁵, Canada (CTaCS)⁶ and the United States of America (USA) (The Urban-Brookings Model)⁷. While models such as TAXBEN are intrinsically static, a further set of models sets actual and hypothetical reforms within a dynamic (lifecycle) framework to analyse changes in policies that are explicitly designed to have effects in the long run. Such models (for example, PENSIM2 in the UK and POLISIM in the USA) can, for instance, analyze how changes in pension rules today affect future recipients of state retirement pensions.⁸

As developing countries attempt to improve public services, consolidate safety nets and redistributive policies (such as Conditional Cash Transfers), there is an obvious need to increase government revenues, through both direct and indirect taxation. This is particularly important in a country such as Mexico: it relies on volatile revenue streams (such as oil) and its revenues from taxation account for less than half of fiscal total revenues. The establishment of new taxes, the increase of the effective tax base as well as the increase in the rate of taxation are all important options that will involve important choices. In making these choices, there is a need to study the distributional impact of tax reform and the implied welfare changes. Distributional and welfare analyses of changes and proposed changes to the tax (and benefit) system are crucial to promote and inform the design of effective and coherent public policy, based on a detailed understanding of the impact that different interventions and changes in the economic environment have on the living standards (and, where possible, the behavior) of individuals and families.

Analysis of the distributional impact of taxation in Mexico will involve a number of practical and conceptual issues that are typically of less concern in a developed economy context. First and foremost, is the high degree of economic informality and tax evasion meaning that the actual tax burden is significantly less than the legislated burden (on average). The issue is more complex than simply assigning someone to the formal or informal sector, as an individual may pay taxes on part of their income (e.g. their earnings from their main employment) whilst not on the other part (e.g. income from informal 'odd jobs'). Moreover, decisions about formality and informality are likely to be affected by the level and type of taxation. Furthermore, the scope for legal tax avoidance is also significant in Mexico and this constrains tax revenues; future work (for instance, making use of taxable income elasticities) will also need to consider this as well as

⁴ Information on IGOTM and PSM is not generally available in the public domain. A simple description can be found in the technical annex to the Lyons review of local taxation (http://www.isitfair.co.uk/Reports/LGR_BOF/final-B.pdf). Talks by government users provide a useful source of information (see http://pcwww.liv.ac.uk/~william/ESRC-BSPS%20MSM%20Seminar%20Series_files/Drane%20transcript%20edited.pdf for PSM, for instance).

⁵ For information on EUROMOD see <http://www.iser.essex.ac.uk/research/euromod>.

⁶ For information on CTaCS see <http://faculty.arts.ubc.ca/kmilligan/ctacs/CTaCS-documentation.2008-1.pdf>.

⁷ For information on the Urban-Brookings model see http://taxpolicycenter.org/UploadedPDF/411136_documentation.pdf.

⁸ For information on PENSIM2 see Emmerson et al (2004) (<http://www.ifs.org.uk/wps/wp0421.pdf>), and for POLISIM see McKay (2003) (https://guard.canberra.edu.au/natsem/conference2003/papers/pdf/mckay_steven-1.pdf).

illegal evasion. A second major concern is the poor quality of income data, particularly for income derived from savings and rents. Solutions to this problem are typically ad-hoc.

In spite of these practical difficulties, there are a number of existing micro-simulation tools for parts of the Mexican tax system. In particular, the United Nations Development Programme is funding Carlos Absalón and Carlos M. Urzúa to develop a micro-simulation tool that can be used by researchers and policy analysts to analyse the distributional impact of actual and counter-factual tax reforms. Like this project, their work also involves building research capacity through a series of descriptive (describing both the tax system and available data) and methodological papers that are accessible to educated non-experts (see Absalón and Urzúa (2009a, 2009b)). Another notable and important contribution to the analysis of tax reform in Mexico is the work of Hector Villarreal and associates (formally working for the Mexican Congress at the *Centro de Estudios de las Finanzas Públicas* (CEPF), and now of the *Centro de Investigacion Economica y Presupuestaria* (CIEP)). Hector Villarreal is a co-author of the present paper, and the analysis and tax simulator models of CIEP have been made available to the researchers engaged in this project. The work of CIEP includes but is not limited to simulators for value added tax, excise duties, income tax due on employment income and social security contributions. A series of papers by Villarreal and co-authors (CEPF (2009a,b,c,d,e)) provides some details on the methods and findings of this work.

We aim to extend and improve upon existing work in two respects. First, we plan an improved consideration of the informal and untaxed sector, further effort to account for missing income, particularly towards the top of the income distribution, and, finally the testing of the sensitivity of our results to different assumptions about these phenomena. Whilst we cannot be sure that such analysis has not been conducted previously, such sensitivity tests have not been published.

Second, our analysis will also consider a number of dimensions of behavioral response to taxation, and we will attempt to analyse how allowing for behavioral response changes the distributional impact of the reforms. Whilst in developed countries, particularly the USA, there has been a significant amount of work investigating the responsiveness of labor supply, saving and consumption to tax policy, there is a paucity of such work for Mexico (although there are ongoing efforts to account for the impact of VAT changes on expenditure patterns, for instance). It is beyond the scope of this study to fill in all the gaps in this knowledge but we will look at two potential responses separately: the choice of goods to consume, and decisions about how much and in which sector to work. It should be noted at the outset that most analysis of the distributional impact of taxation in developed economies abstracts from the behavioral impact of policy changes despite the much better data and existing behavioral models developed for these countries. This means that the behavioral modelling included in this analysis will necessarily be a tentative first step as opposed to the final word on the subject.

The rest of this paper proceeds as follows. Section 2 highlights a number of issues affecting the micro-simulation of tax and welfare changes such as tax incidence, whether to account for dynamic considerations (such as saving) and behavioral response, and issues of particular relevance for developing countries such as informality and non-compliance, and poor quality data. It also includes a discussion of how the distributional impacts of tax and benefit changes are modelled in the UK by the IFS. Section 3 develops our approach for Mexico and includes details on data sources and issues, our tax simulator MEXTAX, together with our chosen approach for behavioral modelling and robustness analysis. It also describes the outputs we

plan to produce and our approach to capacity-building in Mexico. Section 4 discusses possible avenues for future refinement and research. Section 5 concludes. Details of the taxes that we propose to model, and the tax structure and rates in the pre-reform and post-reform systems can be found in the appendices, together with a description of our input data creation procedures and some descriptive statistics.

2. Issues in tax micro-simulation

There are a number of issues to be considered when micro-simulating the distributional impact of changes to the tax and benefit system. In this section we consider 5 of the main issues – incidence, behavioral responses, dynamic factors, informality and non-compliance, and missing or poor quality data – before illustrating how the IFS conducts distributional analysis for the UK in the light of these concerns, drawing on publications produced in the run up to the UK's 2010 General Election.

2.1 Incidence assumptions

In order to characterize the distributional impact of a tax, it is first necessary to ascertain (or assume) the economic incidence of the tax – that is, who bears the economic burden of it. This is not necessarily the same as the statutory incidence of the tax (that is who is legally obliged to remit or pay the tax). For instance, the statutory burden of National Insurance (a social-security payroll tax in the UK) is split between the employer and the employee. However, both portions of the tax have the same economic structure, and hence both parts must, at least in the medium and longer term, be incident on the same agent. In other words, what matters is the existence of a wedge between the cost of an employee to the employer and the employee's remuneration. Who is legally responsible for the various fractions of the payment is not what determines economic outcomes.

Economic incidence of taxation differs from the statutory incidence because of changes in behavior and consequent changes in equilibrium prices. Consider, for instance, an increase in income tax on labor earnings. If the worker's labor supply is not very elastic (i.e. they are unresponsive to changes in their net wages), whilst the firm's labor demand is fairly elastic, the gross wage paid by the firm will rise a little, and the net wage received by the worker will fall a lot. That is, the tax will be largely incident on the worker. If on the other hand, labor supply is highly elastic (i.e. they are highly responsive to changes in their net wages), and labor demand is inelastic, the tax will be mainly borne by firms (the owners of the firm's capital or the firm's customers) in the form of higher gross wages.⁹ So, when considering the impact only through the market for the good taxed (i.e. a partial equilibrium analysis), the relative elasticities of supply and demand determine the economic incidence of a tax. There are further complexities when considering the impact on the wider economy (i.e. a general equilibrium analysis).

Economists have conducted a large number of studies of the incidence of taxation, with most of the research focussing on income and payroll taxes, corporation taxes, property taxes and value added (or sales) taxes. Fullerton and Metcalf (2002) provides a thorough review of the theoretical and econometric aspects of this research, as well as a summary of relevant empirical

⁹ In a similar manner, an increase in the rate of VAT may be borne by consumers (in the form of higher post-VAT prices) or firms and, ultimately, their employees and owners (in the form of lower pre-VAT prices, depending on how elastic consumer demand for goods subject to VAT is).

results. Whilst different studies have found somewhat different results, Fullerton and Metcalf suggest that a consensus has emerged that: payroll taxes are largely incident on workers (Gruber (1997)) as are corporate income taxes (Arulampalam et al (2009)); value-added taxes are largely incident on consumers through higher retail prices (Blundell (2009)); and property taxes are largely incident on existing owners of property as the value of the tax is capitalised.

In most analyses of the distributional impact of tax changes, however, it is not possible to allocate the losses resulting from taxation to different households for all taxes even if one makes specific assumptions about the factor of production or which side of the market the tax is incident on. This is because the survey data typically used does not contain enough information to allocate all taxes (for instance, corporation tax or business property tax) to specific households. Hence such taxes are either left out of the quantitative analysis, or allocated using an approximate ad-hoc procedure (e.g. allocating taxes according to gross income). In the first case it is important to highlight the omitted taxes (and to discuss in a qualitative manner the likely distributional impact of these taxes); whilst in the second case the use of alternative assumptions as a form of sensitivity analysis is advisable.

2.2 Behavioral response

As discussed above, the economic incidence of taxation depends upon the extent to which different agents change their behavior and the consequent changes in equilibrium prices. Apart from this, knowledge of how behavior will change in response to changes in taxation is important to estimate the effect of the policy change on tax revenues, and on economic welfare since those more able to change their behavior are likely to be less affected by an increase in tax. Behavioral response may also be of interest in its own right (e.g. if one is interested in changes in expenditure patterns following changes in commodity taxes).

A full treatment of behavioral response would require a tax micro-simulation tool that was integrated with a full general equilibrium model allowing changes in individuals' labor supply, consumption, saving, in firms' behavior, and in external balance. To our knowledge, such a model is not available for any country – although general equilibrium models with stylised tax systems or representative households do exist (see Plumb (2001), and Schaefer and Peichl (2006), for instance) – and the theoretical, econometric and practical issues involved in developing such a model are very significant (see Adam and Bozio (2009) for a discussion of the difficulties of fully accounting for behavioral response in policy costing – and doing it for distributional analysis would be even more difficult).

However, one can potentially analyse the importance of a number of key dimensions of behavioral response separately using simplified models (subject to the available data being suitable for this task). For instance, the response of consumers to changes in the rate of VAT or excise duties can be estimated using a demand system. Depending upon the choice of demand system it is possible to estimate the welfare impact of the tax changes accounting for the ability to substitute between different goods (e.g. from those subject to the tax increases to those not). In a similar manner, the decision of whether and how much to work, and whether to work formally or informally, can be estimated using a labor supply model.

In some cases, however, it is not feasible to estimate a model of behavioral response. For instance, it is not possible to estimate a sensible demand system if the data on consumer expenditures lacks price variation, and it is often unwise to estimate a labor supply model on a

single cross-section of data (Meghir and Phillips (2010)). In this instance it may be more appropriate to use the existing literature to make assumptions about the relevant behavioral parameters (e.g. the labor supply elasticity), and see how changes in these assumptions affect the behavioral and distributional impact of the tax changes in question.

2.3 Dynamic considerations

The discussion so far has considered the impact of tax changes in a static context – that is, the impact on contemporaneous income and behavior. However, tax changes may also have longer-run effects through, for instance, changes in the incentive to invest in physical and human capital (such as education). Consider the introduction of a new higher top rate of income tax. This reduces income in the short-run, and could lead to changes in the amount those facing the tax will work: for instance some might work more in order to help preserve their after tax income, whereas others might choose to retire sooner if they decide that paid work is no longer sufficiently worthwhile. However, in the longer run it may also reduce the amount people invest in their education and skills because the expected after-tax financial return to those skills are reduced by the higher rate of tax. A number of papers have analysed the long-run impact of taxes on educational choice (for instance, Adda et al 2010), but these models have not been used in the context of estimating the distributional impact of tax and benefit reforms.

Another dynamic concern – borrowing and saving in order to smooth consumption in the face of income shocks – can and should be taken into account when analysing the distributional impact of tax changes. Many people with temporarily low incomes actually have rather higher long-run incomes. An indication of this in survey data is those who have higher spending than their current income. Calculating the proportional loss due to indirect tax changes by dividing the change in the indirect tax by current income can therefore make the reform look much more regressive than it actually would be if it were considered with respect to lifetime income. Under the assumption that the measure of current spending recorded in data is a better guide to lifetime resources than the measure of current income observed in data, the distributional impact of changes in indirect taxes (and indeed other changes) can be better assessed by dividing the change in the taxes paid by expenditure, and using the expenditure as opposed to the income distribution. More generally, the analysis of expenditure inequality and how it is related to income inequality can shed light on the ability of individuals to smooth out shocks and the importance of temporary and persistent shocks (see Blundell and Preston (1998) and Blundell, Pistaferri and Preston (2007)).

2.4 Informality

The term ‘informal economy’ refers to that economic activity that is neither taxed nor comprehensively monitored by the Government and those individuals engaged in such activities. All countries have an informal economy, but the sector accounts for a significantly larger proportion of overall economic activity in most developing countries than in most developed countries. A number of papers estimate that Mexico’s informal economy was equivalent to around 60% of the size of the formal economy in the late 1980s and early 1990s, falling to around 30% of the size of the formal economy since the late 1990s. This is compared with around 8% to 13% in the UK and the USA (see, for instance, Schneider and Enste (2000), Schneider (2002), Brambila-Macias (2008) and Vuletin (2008)). When looking at the proportion

of workers that work informally, Mexico ranks higher than, for example, Brazil and Argentina; other big economies in the region (see, for instance, Galiani and Weinschbaum (2007)).

The presence of informality presents a number of problems for the distributional analysis of tax reform. In particular, by definition, informal activities are not reported to the Government for the purposes of tax collection and therefore no taxes are raised directly from them. However, it is unclear the extent to which the income and spending reported in household surveys picks up the informal as well as formal activities of the sampled households. Ideally, informal activities should be picked up; included in the calculation of a household's total income and expenditure and their place in the distribution of households; and ignored for the purpose of calculating tax payments. However, if informal earnings or spending are not picked up, then one must attempt to account for this when ascertaining a household's total income and expenditure and their place in the distribution of households. It seems likely that survey evidence accounts for some but not all informal activity (in the same way it accounts for some but not all formal activity – see sections 2.5 and 3.2). In a country with a large informal economy, such as Mexico, an attempt to account for informal income and expenditure is therefore both important and challenging.

The potential to work in the informal economy also provides an additional dimension to behavioral response that is of less importance in developed economies: a change in tax or benefit policy may induce not only a change in labor supply but also a change of sector (from the formal to the informal or vice versa). Modelling how individuals and firms react along this margin is key to a full understanding of the effects of tax reform.

A number of papers have set out to analyse informality in Mexico (and Latin America more generally (see for instance Bosch and Maloney (2010))), with some focus on the incentive to engage in informal as opposed to formal work. Meghir et al (2008) estimates a general equilibrium model of the decision to work in the formal or informal sector in Brazil using panel data on employment history and incomes, and administrative data on enforcement of labor market and tax regulations which provides the exogenous source of variation in incentives to formality. In the same way that estimation of consumer demand models or labor supply models requires exogenous variation in prices and wages, estimation of a model of informality allowing causal interpretation of parameters requires such exogenous variation in incentives. This will limit the ability to conduct such analysis, particularly without access to administrative or panel data, and in a short time period (such as available for this project). In section 3.4 we discuss these issues in more detail and set out the main ideas for what will probably be part of a future research project.

2.5 Data quality and availability

The reliability and accuracy of results of microsimulation models depends to a significant extent on the quality of the household data used in the analysis. Ideally, a household survey would be combined with administrative records of benefit receipts and tax payments to provide accurate information on gross earnings, tax deductions claimed for, and benefit take-up (as well as entitlement). However, such data is not available to researchers in most countries due to governmental reluctance to share household-level administrative data (for instance, because of privacy concerns).

Poor quality data can take a number of forms, and in this section we address three main types: (uncorrected) non-random non-response; bias in reported values; and unbiased reported values subject to classical measurement error.

Non-random non-response means that the achieved sample is not representative of the true population. Whilst the construction of sampling weights can partially correct this by re-weighting the data so that household characteristics match administrative and/or census statistics, one cannot correct for non-response that is correlated with unobserved characteristics (or more correctly, characteristics for which external data does not exist, such as the statistics for which household surveys are designed to measure such as household net income, or consumption patterns). Suppose, for instance, that for whatever reason, high income households are less likely to respond to the survey than poorer ones, and that weights can be calculated that re-weight the data so that it matches external figures for various demographic factors (e.g. age, sex, region, tenure, household type etc). The re-weighting can correct for the differential non-response only to the extent that household income varies with these demographic characteristics. Analysis of such data would under-estimate average income, overstate the rate of relative poverty and lead to biased estimates of inequality (although the direction of the bias is uncertain). A tax micro-simulator based on such data would obtain biased estimates of the costs and numbers of people affected by policies that impact differentially across the income distribution. Whilst it is unlikely that such problems will change qualitative conclusions about the regressivity or progressivity of tax measures, quantitative results can be sensitive to non-random non-response (Martini and Trivellato (2003)). In the Mexican context, there is significant concern that the main household income and expenditure survey ENIGH does not cover high income households (see, for instance, Lopez-Calva et al (2007, 2008)). This makes simulation of tax changes affecting this group (e.g. the increase in the top rate of income tax from 28% to 30%) particularly difficult and potentially sensitive to the ad-hoc procedures used to capture the missing income associated with the absent high-income households.

Many surveys find that unearned income is under-recorded compared to administrative data and National Accounts suggesting that higher income households are likely to be under-represented in sample data, even after weighting. However, in the UK, benefits receipts are also under-recorded. Given that benefit recipients are generally low-income (who are, if anything, over-represented in the sample), this suggests that, as well as a biased sample, there is a bias towards under-reporting at least some sources of their income amongst those surveyed. This type of bias has the same kind of effects on analysis of the income distribution as an unrepresentative sample. For tax and benefit micro-simulation, it creates additional problems. For instance, the response to categorical variables (e.g. in the UK context, whether one is contracted-in or contracted-out of the state second pension) and the amounts of certain benefits (e.g. disability benefits) may be used to determine the social security schedule faced by an individual or eligibility for different rates of benefits. Calculating the impact of reforms to these rules, schedules or rates on households would therefore be based on biased data, likely leading to biased results, even for the achieved sample.

If reported values are not biased, but are instead subject to classical measurement error only, problems are less serious but not entirely absent. Because classical measurement error does not affect mean-values, in large datasets, it will not cause biases in the estimated revenues or costs of fully proportional or lump-sum taxes or transfers, nor in their distribution across household

characteristics. However, classical measurement error does lead to an over-estimate of the dispersion of incomes. This means estimated poverty rates and summary statistics of inequality will be upwardly biased. It also means that estimates of revenues from taxes or benefits (and changes in these) that are progressive or regressive will be biased, and that summary statistics of the distributional analysis of such taxes or benefits (and changes) will also be biased. The direction of the bias will depend upon the particular change in consideration. Survey data is always subject to some measurement error meaning that using arithmetic micro-simulation tools to analyse policies affecting different parts of the income distribution differentially cannot provide completely unbiased results, let alone truly accurate results.

All of these problems also affect surveys in developed countries (although, perhaps to a lesser extent than those in developing countries due to greater experience and additional financial resources). Furthermore, in spite these issues, quantitative analysis of tax reforms and the income distribution using survey data remains informative and useful, even if not infallible.

In order to minimise the problems associated with missing households, incomes and measurement error, administrative data can be used to supplement the survey responses. This administrative data can take the form of household-level data where the benefit and tax records of surveyed households are matched with the survey data and are used in place of the reported values. For instance, in the UK, the Department for Work and Pensions is undertaking analysis on linking administrative records with the Family Resource Survey to see how this affects both descriptive statistics (such as average household income, poverty rates and inequality summary statistics), and, potentially, micro-simulation model results. Such data, unfortunately, is often unavailable to researchers outside of government. A second best is to use tabular data on total incomes by various tax-brackets or on expenditures on benefits by household characteristic, to adjust figures so that for these broad demographic or income groups, the survey data and simulator results match administrative figures. If neither of these is possible, an approach based on testing the sensitivity of results to various assumptions about how to allocate missing income (or subtract excess income) is sensible. Adjustment of existing reported values by a constant factor is, on its own, not entirely satisfactory.

A further issue of particular relevance to this project is the type of income data available in the survey. Ideally a survey will contain both gross and net income by various income sources; the former can be used as the basis for simulation, whilst the latter can be used for descriptive analysis and to account for income from parts of the system one does not wish to simulate in the model. If only net income by source is provided, one must first 'invert' the tax system and use this to convert net to gross income. If one wishes to model only changes on tax rates and thresholds all one requires is to convert total net income to total gross income. If one also wishes to model changes in allowances and disregards for particular sources of income (such as those that exist in Mexico for overtime, bonuses, pensions etc.), one needs to convert from net to gross income for each source. As well as requiring knowledge of tax rates and thresholds, this also requires one to know the order in which income is taxed. In order to make a tax simulator as flexible as possible, the creation of disaggregated gross income is advisable.

2.6 The IFS's distributional analysis of UK tax and benefit reforms

Following each Budget and Pre-Budget Report, and at various other times (for instance, prior to General Elections), the IFS produces a distributional analysis of proposed tax and benefit

reforms (or, in some instances, previously enacted reforms). These analyses are high-profile and receive significant media coverage, and help ensure that the Government is held to account for the financial impact of its tax and benefit changes on different groups in society. Analysis is carried out at the family level (or the household level if there are changes to indirect taxation). The analysis is based on the following incidence assumptions:

- Income tax is incident on the individual formally paying the tax (e.g. the worker, or the saver).
- National Insurance (social security) contributions are incident on the worker.
- Value Added Tax and excise duties are incident on the consumer.
- Council Tax (property tax) is incident on the household formally paying the tax.
- Tax credits and benefit payments are incident on the recipient.

For other taxes, such as those not directly attributable to individual households (such as Corporation Tax, Climate Change Levy), or for which the data does not record sufficient information to calculate tax liability (e.g. Capital Gains Tax), the IFS makes ad-hoc assumptions that are clearly stated. For instance, in its analysis of the distributional impact of tax and benefit changes under the last Labour government (see Browne and Phillips (2010)), IFS researchers assumed that most tax and benefit changes that could not be directly modelled had an impact equal to the same proportion of net household income for each household (which might correspond to an assumption of incidence on consumer prices, for instance). In cases where the majority of the gross fiscal 'giveaway' or 'takeaway' would need to be modelled via ad-hoc assumptions, researchers do not analyse the overall distributional impact in a quantitative manner, but instead provide a qualitative assessment (although quantitative analyses may be conducted for the set of reforms that can be accurately modelled). The analysis of the policy proposals in the 2010 UK General Election (Adam et al (2010)) are an example of this.

IFS analysis of the distributional impact of taxation generally assumes no behavioral response (although for some specific changes the likely behavioral response is discussed or estimated, for instance for changes to the top rate of Income Tax). This is because, as discussed earlier, it is not currently feasible to integrate a detailed tax and benefit simulator with a full general equilibrium model of the economy which would be required to calculate the revenue, welfare and behavioral impacts of policy changes. Without such a model, it seems more transparent and consistent to assume no behavioral response than to include only specific examples of response unless this behavioral response is likely to be of overwhelming importance. Calculating the monetary cost of a change in taxes before behavioral response rather than after it is important because changes in behavior themselves entail a welfare cost (that at the margin must be equal to the monetary cost of paying additional tax). Furthermore, calculating the distributional impact under the assumption of no behavioral impact answers an interesting question in its own right: how would people be affected based on their current behavior?

TAXBEN, the IFS's tax and benefit micro-simulator, is a static micro-simulation tool, but in analysis of indirect tax changes, gains or losses are presented as a proportion of both income and expenditure, using both the income and expenditure distributions to classify households as rich or poor. Such analysis has been important in demonstrating that while increases in the main rate of VAT are mildly regressive with respect to current income they are mildly progressive with respect to spending suggesting that the reform is also likely to be progressive

with respect to lifetime incomes (which is what common-sense would suggest given that most items not subject to the main rate of VAT are necessities rather than luxuries).

3 Tax micro-simulation in Mexico

This section sets out our approach to the simulation of tax reforms in Mexico, taking into account the issues and problems identified in the previous section, and the experience of both the IFS in the UK and other practitioners who have undertaken distributional analysis of tax reforms in Mexico. First we review the data sources available and existing work, before describing our general approach to tax and benefit simulation. We then discuss in more detail the organisation and design of our tax simulator, the behavioral modelling to be implemented, and our approach to robustness analysis. We also detail the outputs to be produced and our approach to capacity building and working in partnership with Mexican academics and practitioners.

3.1 Previous efforts at tax micro-simulation in Mexico

As detailed in the introduction, a number of researchers have developed tax micro-simulation tools for Mexico largely in response to a growing desire to produce distributional analyses of tax proposals, such as the 2009 proposals to be analysed as part of this project. The UNDP has commissioned Carlos Absalón and Carlos Urzúa to develop a tax and benefit micro-simulator that could be used to analyse the distributional impact of tax and benefit changes, to produce non-technical documentation for this and to produce descriptions of the tax and welfare system that are accessible to non-experts (Absalón and Urzúa (2009a, 2009b)). A team led by Hector Villarreal, who was employed at CEFP (part of the Mexican Congress) until March 2010, produced similar models and analysis that has been made fully available to the public (CEFP (2009a, 2009b, 2009c, 2009d, 2009e)). This work has been taken forward by Villarreal's new research institute CIEP, who will produce online versions of their simulator so that the general public can conduct simple distributional analyses of a subset of tax reforms.¹⁰

CEFP (2009d) shows that the set of reforms to indirect taxes initially proposed by the government would have led to the poorest 10% of the population (as measured by their position in the income distribution) contributing 3% of the additional indirect taxes payments, whereas the richest 10% of the population would have contributed 29.8% (see Cuadro 1, under 'Iniciativa'). The corresponding figures for the approved package are 2.8% and 30.9% (Cuadro 1, under 'Aprobado ambas Cámaras'), suggesting that the approved package is more progressive than the proposed one. CEFP (2009e) shows that as a result of the proposed and approved change to the maximum marginal income tax rate, the contribution of the poorest 10% of households to the increase in tax payments is zero, whereas the richest households contribute to 67.6% of the total increase in tax payments (see Cuadro 2).

Both Absalón and Urzúa, and CEFP/CIEP assume that income tax is borne by the individual taxpayer, VAT and excise duties by the consumer, and employees' social security contributions by

¹⁰ The Finance Ministry is obliged to produce a distributional analysis of the approved tax changes and includes personal labor income tax, VAT and excise duties. Despite access to official tax-payer data, this analysis is considered to be of only limited use given a reliance on relatively crude assumptions.

the employee. The exclusion of employers' social security contributions from the tax simulators means that at present the distributional analysis of changes to these taxes could not be modelled. Ideally, a simulator should include both employee and employer social security contributions as, in the medium to long run, the incidence must be the same.

Before understanding the approach of previous researchers (which is a necessary first step in defining the contribution of our own analysis), it is important to understand the basic input that all tax micro-simulators require – cross-sectional (or panel) data on households' demographics, incomes and expenditures. In the case of Mexico, the main survey used for the purpose of analysing household incomes and expenditures is the *Encuesta Nacional de Ingresos y Gastos de los Hogares* (ENIGH). This is a biennial survey of households in Mexico and includes 29,468 households in the most recent year of data, 2008. Because of the way the sample is determined (it is not a fully random sample because of a desire to over-represent certain types of households for which there is particular interest, such as rural households), it is important to use sampling weights to ensure the data are representative of Mexico as a whole (although see Section 2.5 for the limits of weighting).

The survey collects information on households and individual demographics, *net* incomes from employment, self-employment, investments, transfers and in-kind sources of income, together with information on a wide variety of expenditures. The fact that the survey captures net as opposed to gross incomes means that, before tax reforms can be analysed, it is important to first use the tax system in place at the time of the survey to 'reverse calculate' the gross income. This can be difficult given the multiple exemptions and allowances that form part of the Mexican tax system (see Appendix A). It is nevertheless an integral part of the programs developed by Absalón and Urzúa (2009b), and Villarreal et al (published by CEFP (2009a)).

A number of other problems have been identified by previous research, with the solutions used not entirely satisfactory. For instance, researchers have found that there is a significant under-reporting of income in the ENIGH survey, particularly for non-labor income. For instance, the Ministry of Finance, using the national accounts for the year 2008 and the ENIGH 2008, calculates that between 16% and 25% of income from employment is not-reported, whilst 92% of non-labor income is not reported (see SHCP (2010)). If left uncorrected, this would affect estimates of the impact of the tax reforms, estimates of average income, and if the degree of under-reporting differs systematically across the income distribution, the characterisation of the income distribution.

Correcting for this under-reporting is not easy. A typical solution is to increase the reported values of labor and non-labor income by the necessary factor for the survey data to replicate national accounts data (see Absalón and Urzúa (2009b) and CEFP (2009a) for instance). However, to the extent that households are engaged in the informal economy, the resulting amounts may remain an under-estimate of total income. Furthermore, it is unlikely that individuals correctly report whether they have labor or non-labor income, but then under-report the amount by a common factor. It seems more likely that, particularly for non-labor income, a large fraction of the under-reporting relates to individuals omitting sources of income completely (or due to the upper part of the income distribution being under-sampled). Indeed it is possible that those reporting non-labor income report the correct amount (or at least an unbiased but noisy estimate of the correct amount), with all of the underreporting being accounted for by those reporting no non-labor income. Unfortunately, there is no clear method

to identify those who completely omit an income source that they receive, although regression-based techniques can be used to assign this income based on the characteristics of those reporting some unearned income in the first place (to illustrate the extent to which assumptions about the nature of under-reporting may drive distributional patterns and revenue estimates).

The ENIGH 2008 survey contains a number of variables that can be used to identify informal activity. For instance, for the purposes of deciding whether someone is formally employed (and therefore paying taxes and social security on their earnings), the survey asks people whether they are covered by the health systems provided to those contributing to the social security system, and whether they are covered through their own work. Unfortunately, for those with two or more jobs, it is not possible to determine how many and for which job those reporting coverage are paying contributions and taxes. It also does not allow us to ascertain whether they are paying taxes on their unearned sources of income. Allocating expenditure to the formal and informal sector can be achieved by reference to the type of vendor the goods were purchased from.

Because variables on social security participation and place of purchase are used to determine whether a transaction is taxed or not, non-random measurement error in these variables could cause significant problems for the analysis. This is something that does not appear to have been addressed by existing work (for instance, by comparing estimates of participation in schemes or place of purchase to other sources of data such as official estimates of evasion, or administrative records on scheme membership).

Existing tax micro-simulation tools have adopted relatively simple methods to address informality and tax evasion. For instance, CEFP analyses consider expenditures at all vendors other than street markets or stalls as formal when calculating indirect tax payments, and there is an assumption that workers defined as formal comply fully with social security and income tax legislation (and others not at all) when calculating social security contributions and income tax payments. Formal workers are defined as those that are covered by the health systems (not necessarily through work). See CEFP (2009a, 2009b, 2009c, 2009d, 2009e).

The Mexican tax collection agency, the *Servicio de Administracion Tributaria* (SAT) estimate that approximately 77% of the Income Tax due on non-salary income is not paid, whilst the evasion rate for Value Added Tax is 35%.¹¹ These figures provide additional information that can be used when estimating tax payments and net incomes accounting for informality and non-compliance, and as a check on the implied evasion rates under the assumptions chosen by the researcher.

3.2 Our proposals

Our proposals draw both upon the experience of the IFS in analysing tax and benefit reforms in the UK, and on the methods used by other researchers in Mexico (particularly the analysis described in the various documents by CEFP (2009a,b,c,d,e) conducted by Villarreal et al at the Mexican Congress and, now, CIEP).

In terms of data, as with previous efforts, our main data source will be the Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2008. ENIGH is a national representative survey of

¹¹ Figures available from the SAT website. An alternate figure of 20% is estimated in Samaniego et al (2006).

around 29,468 households in both urban and rural areas, conducted by the Mexican statistical office (INEGI). It provides detailed information on the households' income and expenditure and other demographic and socio-economic variables.

Because of the poor quality of data on income (unearned income in particular), an adjustment to the raw data in ENIGH 2008 is required. Like previous studies, we will make the adjustments, in the first instance, by increasing reported amounts by a common factor so that the aggregate incomes match national accounts figures. However, we also plan to use other methods, such as allocating additional unearned income based on individual and household characteristics using a regression-based framework, but with an element of randomness (both in terms of amounts, and in terms of who it is allocated to). It should be noted that the use of these alternate methods will not represent an answer to how missing income should be modelled, but will indicate the sensitivity of results to such assumptions (see section 3.5). Ideally, we would also try to obtain information from administrative or other sources on the distribution of unearned income across the population to inform the income adjustments.

There is also a concern that ENIGH sample does not include those Mexican households with the highest incomes (See, for instance, Lopez-Calva et al (2007, 2008)). Administrative data (such as tax-payer records or the Census) would again potentially prove useful in estimating the magnitude of this under-representation and in implementing a solution. For instance one could increase incomes by a factor that varies across the income distribution in a way that makes the distribution of taxable income in the ENIGH sample match that in administrative data. If these data are not available, a sensitivity analysis could again be conducted by making a series of alternate assumptions about the incomes of those high-income households omitted (i.e. different fractions -e.g. proportional to income, half, three-quarters or all- of the income missing from the survey could be allocated to the top decile, for instance).

To account for informality and non-compliance, we will first use the variables in the ENIGH that identify whether someone is covered by the social security health systems (for social security and income tax on earned income), and on the type of vendor from which goods and services were purchased (for indirect taxes). We will then compare the proportion of taxable income and expenditure that is estimated as taxed using this methodology with the figures for tax evasion reported by SAT. As there is no indicator for tax compliance for unearned income we will apply directly the official estimates of non-compliance. Again, the sensitivity of results to the specifics of these assumptions will be tested.

Our treatment of informality and non-compliance must be integrated with our treatment of missing income. This is because, the administrative data that will be used, in part, to account for under-reporting of income in the ENIGH survey itself will suffer from under-recording due to tax evasion.¹² Incomes will therefore be adjusted to account for under-recording of taxable income, and under-recording of untaxed income in a coherent way.

Incidence assumptions will be relatively simple in our main simulations. Indirect taxes will be incident on consumers, and income tax and social security contributions (both employee and

¹² A large fraction on untaxed income, particularly towards the top of the income distribution, seems to be due to the use of tax rules to engage in tax avoidance as opposed to illegal tax evasion, and so will be picked up in administrative records and National Accounts. See "Medición de la Evasión Fiscal en México", coordinated by Ricardo Samaniego, CAEPP, ITAM (2006)

employer) will be incident on the worker. We will account for the fact that income may not be the best measure of long-run living standards (due to borrowing and savings to smooth consumption), by presenting the impact of tax changes both as a proportion of income and expenditure, and by breaking the population into decile groups based both on their place in the income and expenditure distributions. This is in-line with analysis conducted by the IFS in the UK (see section 2.6), and is important to avoid inaccurate and misleading conclusions (such as increases in VAT/IVA being 'regressive').

Mexican researchers and practitioners are particularly interested in how changing the assumption of full pass-through of changes in VAT and duties to consumer prices affects the distributional (and behavioral) impact of changes in these taxes. We plan to investigate the sensitivity of the distributional analysis to a small set of different assumptions about pass-through, drawing on the international literature on VAT pass through (see Blundell (2009), for instance).

In addition to the ENIGH data we may also use the Mexican Labor Force Survey, ENEU (*Encuesta Nacional de Empleo Urbano*), which focuses on 16 urban centres, for the analysis of informality (see, for instance, Maloney (1999)) and its newer version since the year 2005, the *Encuesta Nacional de Ocupación y Empleo*¹³ (ENOE), which includes non-urban areas, as this survey has a longitudinal element. The labor force surveys are quarterly surveys and work so that one-fifth of the sample is removed from the sample each quarter (and replaced by a new set of households) so that each household spends 5 quarters in the survey (unless they are subject to attrition). By observing individuals for five quarters it is possible to construct patterns of mobility between the formal and informal labor markets/sectors.

The main advantage of the ENEU and ENOE surveys is their size and the greater information on labor supply. The main disadvantage is the complete absence of information on consumption and expenditure. An interesting methodological issue is the extent to which the variables that are observed in the two surveys are comparable and the extent to which one can think of integrating the two surveys using econometric techniques for the use of complementary data sources. We plan to explore these issues (A similar technique having been used in the UK context by IFS researchers (see Brewer et al (2010)).

The rest of section 3 discusses our proposals in more detail, including our tax microsimulation model, and our approach to behavioral response and robustness analysis.

3.3 The Tax Microsimulator

In order to estimate the distributional effects of the income tax and indirect tax reforms proposed and enacted in 2010, and to facilitate the analysis of future and counterfactual reforms to these taxes (and social security contributions), we need a tax and benefit microsimulation program. This Mexican Tax Simulator (or MEXTAX for short) will be written and run using STATA, a statistical software package that is widely used by academics and practitioners. In many respects, STATA is not an obvious choice for a tax and benefit simulator. For instance, if more time and resources were available, one might choose a programming language that allows the inclusion of a graphical user interface that could be used for basic tax reforms (for instance,

¹³ See <http://www.inegi.org.mx/inegi/default.aspx?s=est&c=10658>, and <http://www.inegi.org.mx/inegi/contenidos/espanol/prensa/comunicados/estrucbol.asp>

Object Pascal or Visual Basic). Furthermore, whilst STATA performs operations for all observations simultaneously, it is sometimes easier to perform tax calculations observation by observation (which can be done in languages like FORTRAN), particularly when de-bugging the model. STATA has been chosen, however, because it is widely used, and setting the parameters of reform systems should therefore be relatively straight-forward. Section 3.3.1 describes the input data to be used by the program, section 3.3.2 describes the proposed structure of the program, and section 3.3.3 describes the output data created by the program. Please note that as the programs are developed, elements of the following are liable to change somewhat.

3.3.1 MEXTAX Input Data

MEXTAX uses a number of input files derived from the ENIGH. Details on how these different datasets are created, and the do-files used to create the datasets (which can be used as templates for writing do files to create data embodying different assumptions) can be found in appendix C. All data will be labelled in file in English, with a Spanish translation provided in the documentation. The input data consists of the following STATA data files:

- The *household* file contains one observation per household, and includes information on the household's location, the date of interview, household characteristics (such as number of people, household type indicator), household weights, equivalence scales etc. It also contains total monthly household expenditure and net income under the baseline system.
- The *consumption* file contains one observation per household and includes information on the household's monthly expenditure by category of good. These categories are, provisionally: A) VAT exempted i) health services, ii) education services, iii) lottery, iv) public transport, v) transfers, vi) other non-food goods and services; B) zero rate VAT i) food, ii) other non-food goods and services, ii) health and medicines; C) VAT taxed i) food, ii) other non-food goods and services, iii) health-related goods and services, iv) petrol, v) telecoms; and D) VAT and duties i) under 14° alcohol, ii) 14°-20° alcohol, iii) over 20° alcohol, iv) beer, v) tobacco. These expenditure groups are far less detailed than in the raw ENIGH but are detailed enough to allow changes to the existing set of indirect taxes, and the introduction of new taxes such as the IETU. It also contains separate categories for informal and formal-sector spending so that the former are not taxed. The expenditure groups will be aggregated for estimation of the demand model (see Section 3.4.1).
- The *adult* file contains one observation per adult and includes adult demographic variables (such as age, sex, education, employment status), and monthly income figures gross by income source. Each income source will be included twice – once for that which tax is assumed to be paid, and once for that which tax is assumed to be evaded (e.g. through working in the informal sector, for instance).

A key aspect of this project is the use of alternative assumptions about informality, evasion and missing data to test the sensitivity of results to these assumptions. Rather than build these assumptions directly into the MEXTAX program (which we want users to have to edit as little as possible), instead, there will be separate copies of the input data embodying the different assumptions (e.g. under-reporting income by a constant factor versus assuming it differs across the income distribution). See appendix C for more details.

3.3.2 MEXTAX Program Structure

MEXTAX is designed to be used by those with an understanding of the basic structure of the Mexican tax system and who have basic competency in STATA. All do files will be thoroughly documented using in-file comments, in both Spanish and English. Use of shorthand code shall be avoided, and any complex loops or conditional statements shall be fully explained. The comments shall be non-technical and accessible to those with a basic understanding of STATA.

The program is made up of several modules, only some of which the end user will need to be familiar with:

- The *Interface* module contains user-edited instructions on: the directories in which the simulation code, input and output data, and parameter code can be found; the name and location of logfiles; the names of input and output datasets; the existing and reform systems to be used in the analysis; and runs the tax simulator. There are options for whether one wants to conduct the behavioral analysis and what assumptions to make in conducting it.
- The *parameter* modules contain the parameters of the tax system. That is the tax rates, the exemptions and allowances applying to income tax, social security, and indirect taxes. There will be modules for actual tax systems (that should not be changed), and user-edited modules for reform and counterfactual tax systems.

In order to produce a reform system, the user would need to copy the code containing the parameters of the existing (or a pre-existing reform system) into a separate do file and make the necessary changes to thresholds, rates or exemptions, and then in the Interface module do-file, change the reform system(s) do-file(s). No changes to the main program code will be required.

The program itself will then use a number of ‘back-end’ modules that perform the necessary calculations to obtain the tax payments of each individual and household, and perform the distributional analysis:

- The *taxbase* module will use the parameters of the tax systems detailed in the selected *parameters* modules to calculate the tax base for income tax and social security contributions.
- The *dirtax* module will use the parameters of the tax system and the tax base to calculate the income tax and social security contributions that each individual makes.
- The *indirtax* module will use the parameters of the tax system together with the household expenditure data to calculate the VAT and excise duties paid by each household.
- The *household* module will aggregate these changes into the total tax payments by household, and will calculate the differences in tax payments between the base and reform system(s) for each tax, and for all taxes considered in total.
- The *distribution* module will calculate the distributional impact of tax reforms, separately by category of reform, as a proportion of income and as a proportion of expenditure. Outputs collapsed decile / household type data files to the output directory.
- The *quaid*s module will conduct welfare analysis of the indirect tax changes accounting for the fact that people can substitute towards goods that are not subject to the tax

increases and towards the informal economy. Will also recall *indirtax* module based on new consumption baskets. Note that the demand model will be estimated in a separate program and this module will include the parameters from this model. Different *quaid*s modules will be constructed for different behavioral assumptions.

- The *labor* module will model the behavioral response in terms of labor supply and shifts between the formal and informal sectors of the labor market. Will also recall *taxbase* and *dirtax* based upon the estimated changed behavior. As with the *quaid*s model, different versions of the *labor* model will be created to test the sensitivity of results to different assumptions about behavioral response.

The program will make heavy use of scalars (e.g. the number of tax rates, thresholds) and vectors (e.g. the tax rates) as this will aid programming by making it easier to use conditional “if then” loops and “do while” loops, and will reduce implementation time.

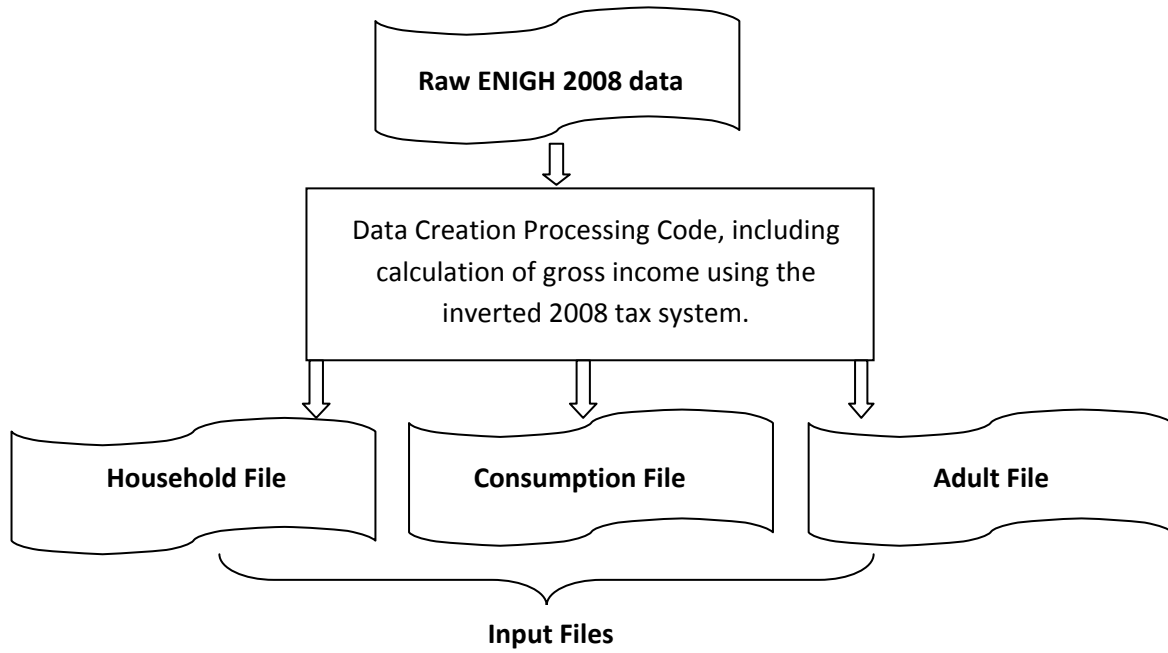
3.3.3 MEXTAX Output Data

MEXTAX creates a number of output files, in STATA format, and in logfile format. On exiting the program, the file loaded into memory will be the *observation output file*. The names and locations of other files will be displayed on-screen so that they can easily be loaded. The files that will be created are:

- The *observation output file* is an adult level file which contains household and adult level demographics, and household and adult level gross incomes, expenditures, and net incomes and tax payments under the various tax systems, together with weights, equivalence-scales, income decile and household type indicators. Given the problems of informality and missing data, the individual data should be used only for conducting analysis at the group-level (for instance, educational attainment or age-group). The individual data itself will be ‘meaningless’ to the extent that the ad-hoc adjustments to incomes will mean that the household data no longer represent individual ‘real’ households.
- The *income distribution file* is a file at the household equivalised decile level, and includes average gross and net incomes, expenditures, tax payments and changes in tax payments (both in absolute and proportional terms) by decile of the income and expenditure distributions.
- The *household-type file* repeats the analysis of the income distribution file, except it is at the household-type level. The same kinds of statistics will be included.
- The *expenditure pattern file* provides summary statistics on how the indirect tax reforms affect expenditure patterns.
- The *labor file* provides summary statistics of labor supply measures.

When running the program using alternate assumptions about informality, missing data etc, as well as changing the input file names (or path) it will be important to change the output file names (or path) to ensure that existing files are not overwritten.

3.3.4 A graphical representation of MEXTAX



PTO

3.4 Behavioral Modelling

As highlighted in section 2.2, it is not currently feasible to estimate a full CGE model that would be required to estimate the full behavioral response to tax reforms and calculate the overall welfare effect allowing for all these feedbacks. Instead, we shall investigate two aspects of behavioral response, separately, to determine how large an impact accounting for these has on the distributional analysis and on tax payments.

3.4.1 Consumer Demand

We aim to estimate the change in the pattern of goods and services purchased following the changes to the tax system using a Quadratic Almost Ideal Demand System (QUAIDS) (Banks, Blundell and Lewbel (1997)). This is a flexible demand system that allows for goods to be luxuries at some levels of expenditure and necessities at others and because it is integrable, one can calculate the exact welfare effects of price (and tax) changes after accounting for the ability to substitute between different goods. Parameters are constant across households but a degree of heterogeneity is possible as the demand system allows for the inclusion of demographic terms in the share equations and price indices.

In order to make the modelling tractable it is necessary to aggregate the goods observed in the survey data into groups with similar characteristics. However, if the aggregated groups are too broad, the model may not capture important substitution opportunities. We need to maintain the number of categories at around 10 to make the model tractable. The groups shall be chosen so that those goods to which VAT and various excise goods and services are applicable are clearly separated from those goods for which these taxes are not applicable. To the extent that taxable goods purchased in the formal and informal sector command different prices, we will consider distinguishing between goods and services purchased in the two sectors. This will be relevant if there are effective substitution opportunities between the two sectors.

Table A.1 in appendix A shows the categories we plan to use in the distributional analysis, which will be the basis for the behavioral analysis. Goods and services have been classified in each tax category using the tax system valid in 2008. These are still more than 10. We will work on reducing them in a way that still allows meaningful substitution patterns for the purpose of the current study. We will also consider disaggregating some of the categories further to define functional groups (e.g. meats, clothing, transportation) if possible.

An initial study of the ENIGH 2008 data reveals that for those goods and services that are subject only to VAT, purchases from informal vendors represent only around 3% of total monetary expenditure reported in the survey (see table A.1), and only for a dozen specific categories is the number of households purchasing from informal vendors significant. For goods that are subject to VAT and duties, only a small number of households (less than 5% of the number of households purchasing these goods) purchase them from informal vendors. For this reason, informal and formal purchases are not enumerated separately for these goods. Table D.3 shows the differences in prices paid in the informal and formal sector for some specific goods that are subject to VAT and for which we observe unit values in ENIGH 2008. Prices paid in the informal sector are statistically lower relative to formal sector prices for food but not for soft drinks.

To estimate the model we need to observe the prices paid for goods and services in each category in 2008; and we require that prices exhibit variation across households (or over time). As a first step, we plan to use detailed data on the prices of foodstuff and other goods and services published by the Mexican Central Bank. The data published online provide price indices for highly detailed goods and services for the 6 broad Mexican regions and for the biggest 46 urban cities in Mexico.¹⁴ In order to construct relative prices in 2008, we plan to obtain data on the underlying series of price levels used to construct the price indices. To use the prices data at the city level, we will use the information on the location of households in ENIGH 2008 to match them to the nearest (or most comparable) city.

We will attempt to construct unit values from data in ENIGH 2008. This is possible for foodstuffs, for which households report the quantity purchased and the amount paid, so that unit values can be constructed for each type of food. For other goods and services it is not obvious how unit values can be obtained. This is because no quantity information is provided. Work done by CIEP researchers has involved constructed unit values for the full set of goods and services in ENIGH 2008 using data on monthly expenditure on each good and a set of assumptions about quantities derived from a careful and extensive analysis of the survey questionnaire and data. It is not obvious that we will be able to do the same work under this project timescale and resources, and it may not be possible to obtain the unit values constructed by CIEP.

There are both positives and negatives about using more aggregate price data than the household-level unit values obtained from ENIGH micro data. Using the unit value paid by each household introduces an endogeneity problem for prices. Variation in unit values may reflect variation in quality as well as in the underlying prices; so the estimated elasticities will reflect changes in behavior related to changes in both prices and quality. A further problem arises if there is non-linear pricing. Take, for example, bulk-purchase discounts. Households purchasing small quantities of a good will pay high prices; and those buying large quantities will pay low prices. In this instance it is quantity that drives price rather than the other way, as would be assumed by the demand model. The estimated elasticity of substitution will be upwardly biased (in absolute magnitude) in this case. Using aggregated price data will overcome these problems.¹⁵

As mentioned above we also plan to investigate the sensitivity of the behavioral analysis to a small set of different assumptions about pass-through, drawing on the international literature on VAT pass through (see Blundell (2009), for instance).

The model shall be estimated using an iterative method outlined in Phillips (2009). Standard errors can be estimated using clustered bootstrap.

3.4.2 Labor Supply and the Informal Sector

Changes in the rates and structures of both direct taxes (such as income tax) and indirect tax (such as VAT) can influence individual's labor supply decisions. By labor supply one means not only the hours of formal work, but also the effort put in during working hours, the choice of

¹⁴ See "Series Índice de Precios al Consumidor - Por objeto del gasto y actividad económica - Por región y Por ciudad" at <http://www.banxico.org.mx/politica-monetaria-e-inflacion/estadisticas/inflacion/indices-precios.html>. Broad regions are Frontera norte, Noroeste, Noreste, Centro norte, Centro sur, and Sur.

¹⁵ Attanasio, Di Maro, Lechene and Phillips (2010).

whether to be employed or self-employed, and whether one works formally or informally. Ideally, this project would consider a number of different margins of adjustment, and a number of different issues. We consider the following issues to be of key importance:

- The impact of the increase in the top marginal rate of income tax on the incentives for high-income individuals to evade and avoid tax, including by increasing the incentive to become self-employed as opposed to employed. Whilst rigidities in employment contracts may limit the ability of employees to adjust their working hours in response to reform, self-employed individuals would not face such issues, and the increase in the top rate of tax could increase incentives to become self-employed to take advantage of this flexibility and use the existing tax loopholes available to this group. A study of the responsiveness of higher-earners to changes in incentives to avoid tax through self-employment would inform estimates of the revenue impacts of changes in the top rates of income tax.
- As well as an explicit study of incentives to self-employment, calculation of taxable income elasticities (à la Feldstein) will provide very useful information on the welfare and revenue impact of changes in tax rates, and provide a useful summary statistic for overall behavioral response.
- How (exogenous) changes in incentives individuals face to work in the informal sector affect whether workers are formally employed or not. For example, individuals may prefer to work in the formal sector if it means they get insurance they would not get in the informal sector. Employers' incentives to employ workers informally may also be affected by reforms, and this should be taken into account. The introduction of Seguro Popular in 2001 (a health insurance provided to low income households not covered by social security) provides a good natural quasi-experiment to assess this type of question. Seguro Popular was first introduced as a pilot in specific states, and was gradually rolled out across the rest of the country.¹⁶ This may provide the necessary exogenous variation in incentives to work informal over time and across regions. Once the importance of these incentives are understood, we could then incorporate this analysis to assess how changes to the financial incentives to work in different sectors affect the revenues achievable from tax changes.
- An exploration of how changes in indirect taxes affect labor supply decisions (recognising that as well as reducing the return to work, indirect taxes may impact labor supply because of non-separabilities between consumption and leisure).

It is our view that these projects cannot be accomplished within the time available for this project, for the reasons detailed in our discussion of the literature and empirical estimation requirements below. Instead we propose to conduct sensitivity analysis using different assumptions about labor supply and taxable income elasticities, drawing on the literature that exists in both developed countries (see Meghir and Phillips (2010)), similar developing countries, and where possible, in Mexico (see Gong and van Soest (2002)). We will also further develop the research ideas outlined above, highlighting the methods that would be informative, the data requirements, and the identification assumptions and strategy proposed for future work.

¹⁶ See, for example, Nigenda (2005) and Scott (2006).

In order to estimate the impact of changes in the tax system on labor supply requires the identification and estimation of a model of labor supply. To be able to use the results of such an estimation exercise to analyse the effect of behavioral response on the welfare impact of tax changes, the model and its parameters have to be 'structural'. By that we mean that the models parameters are invariant to policy changes. This is the case if the estimated model is based on an explicit utility-maximization problem as opposed to consisting of a reduced form labor supply equation. Explicit modelling of labor supply decisions, often involving discrete decisions (e.g. to work or not work, or choice of sector), is not easy and presents considerable challenges, especially if one wants to take into account dynamic considerations, such as returns to experience.

Whilst the distributional analysis of most tax and benefit policies abstracts from behavioral responses, there is a sizeable and growing literature that investigates the responsiveness of labor supply to policy changes, and can be used to assess how these behavioral responses affect government revenue, and household incomes.¹⁷ In order for inferences based on these models to be robust, the models must be estimated using data that contains plausibly exogenous changes in work incentives (for instance, due to changes in policy, policy differences across households or individuals who would otherwise be expected to behave similarly, or secular changes in wages). This generally requires the use of panel data or repeated cross-sections so that one can control for unobserved factors that influence both hours and wages and that would otherwise lead to biased estimates (Meghir and Phillips (2010)). Estimation of such a model for Mexico would require the use of multiple years of a surveys with full income and demographic information (such as ENIGH), the identification of the natural experiment which would form the basis for model identification (such as reforms to marginal tax rates), and the design and implementation of a method to properly control for the other characteristics that influence hours and wages. Our view is that this is not feasible given both the short timescale of the current project, and the significant problems of measurement error in the ENIGH, the main survey that will be used in this project. We will however investigate the feasibility of using alternative data sources (such as ENEU/ENOE) for such a purpose.

A further complexity when analysing the impact of taxation on labor supply in Mexico and other developing countries is the existence of a large informal sector. A significant literature has developed characterising and seeking to explain the nature of informal employment in developing countries. Recent work has questioned the traditional view of informal sector work as necessarily inferior to formal work, operating as a sector where those unable to obtain formal sector work subsist (see Maloney (1999), Maloney (2004) and Bosch and Maloney (2010)). Instead this work finds that there is significant movement in both directions between formal and informal work, and that a significant number of workers in the informal sector are in that sector through choice. Whilst this suggests that moves from formal to informal work need

¹⁷ van Soest (1995) is an early example of a discrete choice model incorporating the tax and benefit system for the Netherlands, and Das and van Soest (2001) use a similar framework to analyse proposed tax changes in that country. These papers do not consider revenue or household income effects, however. Peichl and Schaefer (2006) describes a German tax simulator that is integrated with a similar labor supply model more fully than in the original papers. Brewer et al (2007), Bell et al (2007), and Meghir and Phillips (2010) combine the IFS's tax and benefit microsimulator for the UK, TAXBEN with discrete choice models of labor supply to analyse the impact of actual and counterfactual policy reforms. Bell et al (2007) explicitly analyse how the costs of reforms and household incomes change as a result of behavioral response.

not necessarily involve significant losses of welfare, it does emphasize that such shifts in sectoral composition could be a significant form of behavioral response to taxation.

A number of papers seek to estimate the impact of changes in labor market regulations and income and payroll taxes on the size of the informal and formal sectors. Albrecht et al (2009) develop a model of the labor market with job search and job matching, allowing for the existence of an informal sector in addition to a formal sector, and for workers to vary in productivity along a continuous distribution. Rather than estimate their model using actual data, they use numerical simulation using synthetic data and in their baseline model calibrate parameters so that the estimated rates of formal and informal employment, and sectoral exit and entry rates, and payroll and severance taxes are consistent with Latin American economies. They simulate the impact of payroll taxes of 0%, 50% and 100%, and find that an increase in payroll tax reduces the proportion of people who would only accept formal employment, and increases the proportion who never accept formal employment, and that employment duration falls. However, if there is severance tax, the payroll tax is welfare-improving (in that it raises net output plus tax revenue) because it reduces distortions due to the severance tax and excessive job search. It should be noted that the changes in payroll taxes considered in this paper are far larger in magnitude than the changes to the top rate of income tax in Mexico. Schneider and Enste (2000) cites research which also identifies high social security contributions and tax rates as an incentive to informality. For instance, Cebula (1997) finds that a 1 % point increase in the US federal income tax rate would increase the size of the informal economy by 1.4 % points (although it should be noted that this seems to be at odds with other evidence such as taxable income elasticities, which would generally imply a much smaller effect).

Other papers find much less of an effect of payroll taxes on the size of the informal economy (see Botero et al (2004), and Ulyssea (2010)). Ulyssea (2010), for instance, uses Brazilian data to estimate the impact of entry costs, enforcement efforts, payroll taxes and income taxes on labor market composition. He finds that changes in payroll taxes have little impact on informality rates – for instance cutting the payroll tax from 35% to 15% would reduce the size of the informal sector by only 2% points in this model.

To the best of our knowledge, models integrating a detailed tax and benefit simulator with a model of the formal and informal labor markets have not been developed in any country. Instead, models based on stylised depictions of the tax system are the norm, meaning that distributional analysis cannot be conducted. It is also important that models of the informality decision be estimated using data that contains exogenous variations in incentives for informality (such as policy changes, differences in enforcement or taxation across regions or markets, etc.) given that other factors are likely to drive both wages and sectoral choice. Again, panel data is likely to be important in overcoming these concerns, with the Mexican ENEU/ENOE labor force surveys potentially useful in this regard. The difficulty is, again, pinpointing a source of exogenous variation in informality incentives that can be used to identify the causal effect of taxes (as opposed to social policy) on informality rates (see above). We feel it is necessary to leave such work to future research projects with sufficient time and resources available.

3.5 Robustness analysis

Because of significant problems in the underlying survey data (for instance, missing income and households), uncertainty about the incidence of taxes (such as VAT), and the difficulties identified in estimating behavioral models in the time available, this project will make significant use of variant assumptions to test the robustness of the analysis and findings. The estimated distributional impact of reforms may be significantly affected by the methods used to account for missing income and expenditure, the definitions of formality and informality used, the equivalence scales used to adjust household income for household size, and the choice of whether to use income or expenditure as the indicator of household resources. We have noted in previous sections a few of the types of robustness and sensitivity analyses we have planned. In this section we present these proposals together and provide some more detail.

It should be noted that we are not attempting to provide a definitive answer for how any of these issues should be modelled. Instead, we wish to show how results change when different assumptions are made about missing income, informality, tax incidence, etc.

3.5.1 Accounting for missing income and expenditure

As detailed in section 2.5, accounting for the under-recording of aggregate income in surveys such as ENIGH by increasing them by a factor so that they match administrative or National Accounts aggregates is not entirely satisfactory. Whilst the linearity of expenditure taxation, the presentation of analysis at the group level (e.g. income decile group), and the use of proportional (as opposed to cash) changes in net income means that constant factors are suitable for expenditure, non-linearities of the direct tax system mean that the use of constant factors can bias results. We therefore plan to test the sensitivity of results to a number of different ways of allocating this missing income:

- Using constant factors as is existing standard practise.
- Using factors that vary (smoothly) across the income distribution to account for the concern that it is mainly towards the top of the income distribution that income is under-reported and households missing
- Using a regression-based approach to allocate missing earned and unearned income based on the characteristics of individuals and households, allowing for a degree of randomness by drawing randomly from the error distribution. Ideally one would like to 'predict' missing income separately for each category in the tax simulator, but it will probably be more feasible to predict for unearned and earned income separately and then allocate to sub-components using either the individual's observed sub-component shares (or some average of the shares across households)

3.5.2 The definition of informality

As detailed in sections 3.1, 3.2 and appendix C, in our baseline survey we will define someone as formal if they are covered by the social security healthcare system through their own work, and where they report coverage under more than one scheme (e.g. both IMSS and ISSSTE) because of coverage due to both one's own work and one's spouse's work, we will use information on their main employer to allocate them to a specific scheme. The work of CEPF/CIEP seems to condition only upon whether one is covered by a social security system, and where one is

covered by both IMSS and ISSSTE, the assumption is that it is ISSSTE contribution that one pays. We will test how important these differences in assumptions are for the distributional analysis.

We also plan to compare the rates of informality implied by our method (and the method of CEPF/CIEP) to estimates of the proportion of Mexican workers employed (or self-employed) informally. Where these are significantly different, we plan to randomly allocate some workers to be formal (or informal) so that we match the make-up of the Mexican workforce, and again, see how important this will be for the distributional analysis.

Tax evasion (and legal avoidance) is higher for non-labor as opposed to labor income. This means it is highly likely that some people correctly report their labor income to the authorities (or it is reported automatically by their employer) whilst evading the taxes due on their self-employment income, business and other sources. Hence, it is not desirable to base a definition of tax evasion for non-labor income on coverage by social security healthcare. If we are able to model taxes on unearned income (as opposed to simply accounting for them as a source of net income), we will therefore test the sensitivity of results to different assumptions about the reporting of non-labor income to tax authorities. The nature of these assumptions will be developed in conjunction with the World Bank and CIEP.

We define expenditure as formal or informal based upon the type of store the good or service was purchased from. Table A1 shows that this leads to an estimate of VAT evasion considerably lower than that estimated by the Mexican government. We therefore plan to test the sensitivity of results to randomly reallocating spending from the formal to informal sector to match estimated evasion rates. We will investigate whether this procedure can differ by good recognising that expenditure on certain goods (e.g. utilities) is unlikely to escape VAT whilst others (e.g. household maintenance) are more likely to escape VAT. This is subject to obtaining information on evasion by type of good and service.

3.5.3 The incidence of VAT and excise duties

As detailed in section 2.1 and 2.6, most micro-simulation models including the IFS's TAXBEN and the proposed MEXTAX allocate the impact of changes in indirect taxes to households based on their purchases. That is, they assume that indirect taxes are incident on consumer prices. Attempts to estimate the part of VAT increases borne by consumers (termed VAT pass-through) have generally found that prices rise to largely but not necessarily fully pass the burden to consumers (Blundell (2009)), with the degree of pass-through increasing in the degree of product-market competition in some studies (Carare and Danninger (2008)). If producers bear part of the impact of increases in excise duties in the form of lower profits or wages, the distributional impact of these taxes may differ and we therefore plan to investigate this proposition.

Drawing on the literature we will choose a small number of alternate assumptions about pass-through (e.g. 50% and 75%), and will allocate gains/losses to labor and capital income using the capital/labor share of income for Mexico. Losses will be proportional to income, and we may conduct the analysis both including and excluding labor income derived from the public and the informal sectors.

To model less-than-full pass through, pre-tax prices will fall by an amount so that for the given increase in the VAT (or excise) rate, the pre-tax price falls so that the increase in the tax

inclusive price is equal to 75% (or 50%) of what it would be under the assumption of full pass-through. The direct impact on consumers can be calculated as the change in post-tax prices, whilst the impact through wages and interest income will be estimated by reducing pre-tax incomes.

It should be noted that this sensitivity analysis will not tell us to what extent VAT is passed through but, instead, what impact less-than-full pass-through could have on the results of distributional analyses.

3.5.4 The choice of equivalence scale

Because households have different sizes, and hence different 'needs', household income (or expenditure) needs to be adjusted for household size when assigning households to the different income/expenditure groups that will form the basis of the distributional analysis. CEPF and CIEP adjust household income for household size by dividing by the number of household members. Households are then assigned their position in the income distribution based on per-capita income. However, to the extent that there are economies of scale within households, and children require fewer resources than adults, this may mean that large households and households containing children are placed lower in the income distribution than they should be. For instance, in EU countries, the modified-OECD equivalence scale is used whereby second and subsequent adults (aged 14 or over) are assumed to require 50% of the resources of the first adult for the household to achieve the same living standard, and children are assumed to require 30% of the resources of the first adult. Some researchers have argued that scale economies are likely to be less important in poorer countries (due to lower housing costs, for instance), and hence we will also consider the case where additional adults require 80% of the resources of the first adult, and children require 55% of the resources. For Mexico we will define adults as those aged 12 or over rather than 14 or over.

Rows 1 and 2 of table D1 in appendix D show the correlation between per-capita income and income using the 50-30, and 80-55 equivalence scales (rows 4 and 5 repeat this analysis for expenditure). They show that, particularly for those households with incomes below the median household income, the choice of equivalence scale may be of some importance. This is because whilst the correlations between the measures are high, they are not perfect.

3.5.5 Income or expenditure as a measure of living standards

Section 2.3 highlighted how saving and dis-saving associated with a desire to smooth consumption in the face of volatile income means that income may not be an appropriate measure of living standards on which to base distributional analysis, particularly for indirect tax changes. This means that whilst when analysing direct tax changes our baseline distributional analysis will be based on a household's position in the income distribution, for indirect tax changes it will be their position in the expenditure distribution. However, for both the analysis of direct and indirect tax changes, and all tax changes taken together, we will calculate gains and losses as a proportion of both expenditure and income, and classify households into decile groups based on both expenditure and income too. Analysis at the IFS has shown that the results one obtains using these different methods can differ significantly, even changing one's view of whether a policy is progressive or regressive (see Mirrlees et al (2010)).

Appendix D contains a number of descriptive statistics about income, and expenditure, and the links between the two.

Figures D.1 and D.2 show the distribution of total income and monetary income, respectively, using a cumulative frequency graph, whilst figures D.3 and D.4 repeat this analysis for expenditure. These show the importance of non-monetary income and non-monetary expenditure across the income distribution (for instance, imputed rent is over half of non-monetary expenditure, on average, and is important for rich as well as poor households). For instance, whilst approximately 35% of households have a per-capita total income of less than 5000 pesos per annum, for monetary income the fraction is approximately 50% of households. Similarly, whilst about 12% of households have a per-capita total-income of more than 20000 pesos; this is true for only 7% of households for monetary income. The figures also demonstrate that there is more variation in income than expenditure, particularly towards the very bottom and the tops of the distribution, as one would expect if households are saving and dis-saving to smooth consumption.

Figure D.5 shows the 25th, 50th and 75th percentiles, as well as the mean of households' positions in the distribution of household expenditure by income percentile. This shows, as one would expect, that households towards the bottom of the income distribution are also towards the bottom of the expenditure distribution, and vice versa. However, the correlation is not perfect.¹⁸ For instance, for households in the middle of the income distribution, whilst the median position in the expenditure distribution is the 49th percentile, 25% of households are in the bottom 34% of the expenditure distribution, 50% in the next 25%, and 25% found in the top 40%. This shows that whether one uses income or expenditure to define living standards can play a significant role in where particular households are in the distribution.

Figure D.6 is similar to D.5, but uses monetary amounts for income and expenditure rather than percentiles. This shows that the poorest 40% of households spend somewhat more than their income, on average, with the richest 5%, and particularly the richest 1% of households spending considerably less than their income. This suggests that there is saving and dis-saving occurring, although measurement error in income and expenditure would also generate such a pattern and is likely to play some role. Rows 7 and 8 in table D.1 show the correlation between total income and total expenditure, and monetary income and monetary expenditure, respectively. These show that the correlation is considerably higher in the lower half of the income distribution, possibly due to credit constraints. This means that the use of expenditure as opposed to income percentiles is likely to have more of an impact when reforms affect richer households more, as opposed to poorer households.

3.6 Structure of the Papers

The final report shall be presented as a suite of three papers (or a single paper with three distinct sections). The first paper will include the basic distributional analysis without accounting for behavioral response. It will detail the impact of tax changes separately for direct and indirect taxes, and will present the distributional analysis both in terms of income and expenditure. The robustness and sensitivity of the results to assumptions about the nature of

¹⁸ It is stronger than the correlation observed in the UK, however, particularly towards the bottom of the income distribution, where many households are much further up the expenditure distribution (see Brewer et al (2009)).

economic informality (for instance, what types of households are most engaged in the informal economy) and the method of adjusting for missing income will be tested. The paper will also include more specific details about the tax simulator MEXTAX and the assumptions underlying the distributional analysis.

Paper 2 will investigate how the changes in the tax system will affect consumption patterns, and how accounting for this behavioral response affects the distributional impact of the reforms, the revenue to be raised etc. The amount of progress that can be achieved on this paper depends on obtaining plausible price variation for Mexico to estimate a demand model.

Paper 3 will include an analysis of the impact labor supply response has on the analysis of the tax reforms. This will make use of representative elasticities and robustness-checking, and subject to feasibility, estimates of the extensive-margin elasticity in Mexico.

3.7 Collaboration between the IFS and ITESM/CIEP

The research produced by this project will be carried out by a close collaboration between IFS researchers, based in London, and Hector Villarreal and his team based in CIEP and ITESM, Mexico. Both teams will benefit from the partnership. IFS researchers have experience in conducting this type of analysis for developed economies, particularly in the UK. Researchers at IFS have strong academic records and experience in translating their rigorous analysis into prompt outputs that are accessible to policy-makers and a wider audience. The Mexican team has a crucial institutional knowledge of the Mexican tax system, the local economy and political economy, and the related work that has been already done in Mexico. It has experience in using the data we will use in our analysis, and has already generated some of the distributional analysis and the codes that are the basis for the current collaborative research project.

We expect that this partnership will build research capacity, both at IFS and CIEP. We also expect that this collaboration goes beyond the limits of this specific project. We anticipate that this will be long-term research collaboration on the modelling of fiscal reform and behavioral response to tax and social policy changes in Mexico and other developing countries.

4 Proposals and scope for future work

This paper has consistently highlighted that the work to be conducted as part of the present project is preliminary and tentative, and that there is much scope for future research that will improve the distributional and wider economic analysis of tax reform in Mexico. Rather than repeat the previous discussions, this section provides a brief summary of the specific recommendations and ideas (with reference to where further details can be found), together with an evaluation of the merits of the development of a fully behavioral microsimulator that takes account of the general equilibrium effects of tax changes, with a more ad-hoc approach based on an arithmetic simulator and specific models for specific types of behavioral response.

Sub-sections 4.1 to 4.4 list areas that we think would benefit from further research and where further research is possible. Subsection 4.5 discusses the merits of an integrated microsimulation tool and general equilibrium model.

4.1 Improving the data on incomes

Sections 2.5, 3.2, and 3.5.1 detail how we plan to test the sensitivity of our results to a number of different assumptions about how the missing labor and non-labor income is distributed across the households surveyed in ENIGH. This includes an attempt to gain access to tabular administrative data which can be used to work out at which parts of the taxable income distribution the ENIGH under (or over) records income so that appropriate adjustments can be made.

It is important to note that none of the assumptions that will be tested represents the best way to improve data on incomes. As a first step, given that ENIGH is being used increasingly for policy analysis as well as more descriptive analysis (such as calculating poverty rates and expenditure patterns), additional effort should be placed on improving coverage of high income households that are currently under-represented in the survey, and in improving the sampling weights as far as possible. Ideally, the government should also link the survey data with administrative data in a similar manner to the Department for Work and Pensions in the UK. However, if permission is required there may be concerns that permission will be non-random (e.g. richer households may be less likely to give permission). Mexican researchers have also suggested using the census (which theoretically covers all households in Mexico and includes questions on income) to adjust ENIGH weights or incomes to account for under (or over) representation of households at different parts of the income distribution.

We think that improvements to the incomes data are important and would welcome the opportunity for future collaboration with the World Bank and Mexican researchers and officials in taking this idea forward.

4.2 Improving the modelling of labor market behavior

Section 3.4.2 details the areas which we think future research on labor market behavior should focus on if it is to be of most use for public policy analysis. To summarise, this research is:

- The impact of income tax and social security on the incentives for individuals to evade and avoid tax, including by increasing the incentive to become self-employed as opposed to employed.
- The calculation of taxable (and 'broad') income elasticities as a summary statistic of behavioral response.
- The use of the introduction of Seguro Popular as an (exogenous) change in incentives that individuals face to work in the informal versus formal sector to investigate how these affect whether workers are formally employed or not.
- An exploration of how changes in indirect taxes affect labor supply decisions (recognising that as well as reducing the return to work, indirect taxes may impact labor supply because of non-separabilities between consumption and leisure).

We also believe that if a discrete choice labor supply model is to be estimated for Mexico (for instance using ENEU/ENOE), it is important that this takes into account sectoral choice (e.g. formal employment, formal self-employment, informal work) as well as hours of work, as this margin of behavioral response is likely to be very important. This will be significantly more demanding than the discrete choice models in use for developed countries which generally abstract from the issue of informal work, and will require close collaboration between

econometricians and those with detailed knowledge of the characteristics of employment (both informal and formal) in Mexico.

4.3 Expanding coverage to include cash welfare transfers

For this project, the MEXTAX model will include VAT, duties, social security and personal income tax (on labor income, and possibly non-labor income). In future it is important to extend the model to include cash welfare transfers such as Oportunidades payments as these are an important source of income for low-income Mexican households, and the Government plans to increase spending on social welfare going forwards. The overall progressivity (or regressivity) of government taxes and transfers is often largely determined by the transfer as opposed to the taxation side.¹⁹

Inclusion of public spending on public services is conceptually and practically more difficult than for cash benefits. In particular, the monetary benefit of the same amount of public spending on a particular in-kind service may differ across households (because of tastes). Because we have little evidence about how willingness to pay for different services varies across the income distribution or across household types, analysis of the distributional impact of public spending on public services often assumes that cash benefits equal cash costs.²⁰ Furthermore, a significant proportion of spending is on public goods (such as Defence), the benefits of which cannot easily be allocated to specific households. Assumptions about how the benefits of such services are shared across the population (such as equally in cash-terms, or equally in proportion to income) can make a significant difference to conclusions. These conceptual and practical difficulties lead us to recommend that, in the short-run, research should focus on the more tractable problem of including cash welfare-payments.

4.4 Making the model more user-friendly

MEXTAX is envisaged as a STATA program. As detailed in Section 3.3, STATA is not the best program for tax and benefit micro-simulation. In future it may be worthwhile developing a program using an alternate programming language more suited to the task (such as Fortran) and integrating the model with a graphical user interface to allow greater ease of use for both researchers and non-expert users of the program. There is a growing movement to make micro-simulation tools available to the general public through the internet, and this would be something which we would support. It should be noted that whilst the IFS manage their own interactive tax and benefit microsimulator, the development of a full graphical user interface would probably be best undertaken by professional programmers as opposed to economists (although it is important that the programmers work in conjunction with economists to ensure the interface allows the types of policy changes and analyses deemed to be important).

¹⁹ See for instance, http://www.statistics.gov.uk/articles/nojournal/Taxes_Benefits_0809.pdf, for the UK. Table 1 shows that the average 'original income' for the top 20% of the population is 14 times than for the bottom 20%, and that after benefits are taken into account this falls to 7 times. Direct taxes and social security payments reduce this only a little more (to 6 times).

²⁰ See presentation by Cormac O'Dea of the IFS in response to the UK's 2010 Spending Review, available at: <http://www.ifs.org.uk/budgets/sr2010/publicservices.ppt>.

4.5 The merits (and demerits) of general equilibrium modelling

Accounting for the full behavioral response to tax changes requires the development of a competitive general equilibrium model that is fully integrated with a micro-simulation model. This would allow not only first round responses to tax changes (such as falls in labor supply following an increase in income tax) but also second round (and further) effects (such as the impact of lower labor supply on consumer expenditure and firm behavior). As detailed in section 2.2, to the authors' knowledge no such models exist, although there are models that include either a stylised representation of the tax system, or are based on representative households. As Adam and Bozio (2009) highlights, to implement such a model perfectly "would require answering almost every question asked in economics. It is not only unfeasible now, it will never be feasible. But perfection is too exacting a standard". Even accepting that our models can only be an approximation (albeit, hopefully an unbiased approximation) to the true answer, they also emphasise that "we have no idea of the magnitudes of likely responses by households and firms to many kinds of tax reforms", even at the Macro-economic level, let alone at the household-level. This does not mean that research on the development of microeconomic general equilibrium models should be abandoned, but we would argue that policy-focussed research should instead emphasise improving partial equilibrium analyses of certain kinds of behavioral response (such as labor supply or consumption).

Models of labor supply or expenditure will be able to take into account more margins of adjustment, more individual heterogeneity, more institutional features, and can be more easily tailored for the analysis of specific policies than a general-purpose general equilibrium model. They can also draw on a larger and more developed literature in developed economies, and are generally less demanding in their data requirements (for instance, not requiring data on firm behavior, or external trade, for instance). Whilst it is laudable to wish to push the boundaries of economic research when analysing tax policy in developing countries, for potentially high-profile distributional analyses it is perhaps more wise to produce and present a simpler analysis that abstracts from some issues or responses, but for which one is more confident.

5 Summary

This paper has set out the main issues in the distributional analysis of tax reforms both generally, and in the Mexican context specifically, and has described and explained our proposed methodology for addressing them. We believe the main issues are the quality of incomes data (particularly for the top of the income distribution), understanding and accounting for the impact of taxes on informality, evasion and avoidance (as well as more traditional labor supply responsiveness), and in ensuring that the tax and benefit simulator is as flexible as possible and can therefore be used to model more than marginal changes to the existing tax and benefit system.

This project is necessarily only the first part of a longer-term research agenda that we hope can make significant progress in addressing the issues highlighted above. Because of the significant time and resources required to link survey data with administrative data or other surveys, and because the development of robust behavioral models requires the identification of exogenous variation in incentives and much refinement, we will instead focus on testing the sensitivity of results to various assumptions about the data and about behavioral response. This will help show how robust conclusions about whether policies are "regressive" or "progressive" are to

deviations from the assumptions used in existing analysis: namely, no behavioral response and income missing in proportion to income reported. Such analysis has not been published for Mexico, and we hope that if these issues are shown to be important, it will act as a spur both to future research and government action to improve the data by allowing use of administrative data, at least at the aggregate tabular level.

This work will build capacity both at the IFS and in Mexico, and it is hoped that this could be the start of a long-term and fruitful working-relationship between the IFS, CIEP and the World Bank. With Mexico facing a significant challenge in filling the fiscal hole that will develop as oil revenues decline, a thorough distributional and behavioral analysis of tax reform will be vitally important in the coming decade, and we hope to be able to contribute to this research agenda beyond this specific project.

References

- Absalón, C. and C. M. Urzúa (2009a), "Mexico Country Report", Fiscal Schemes for Inclusive Development, United Nations Development Program
- Absalón, C. and C. M. Urzúa (2009b), "Notes on the FDI PIT-Simulator", Mimeo.
- Adam, S. and A. Bozio (2009), "Dynamic Scoring", OECD Journal on Budgeting, Volume 2009/2, Paris: OECD. Available at: http://www.ifs.org.uk/docs/dynamic_scoring.pdf.
- Adam, S., M. Brewer, J. Browne and D. Phillips (2010), "Tax and benefits: the parties' plans", Institute for Fiscal Studies Election Briefing Note 13, London: Institute for Fiscal Studies, available at: <http://www.ifs.org.uk/bns/bn100.pdf>.
- Adda, J., C. Dustmann, C. Meghir and J. Robin (2010), "Career progression and formal versus on-the-job training", Institute for Fiscal Studies Working Paper 10/13, London: Institute for Fiscal Studies, available at: <http://www.ifs.org.uk/wps/wp1013.pdf>.
- Albrecht, J., L. Navarro and S. Vroman (2009), "The Effects of Labour Market Policies in an Economy with an Informal Sector", The Economic Journal, Vol. 119., No. 539, pp. 1105 - 1129
- Arulampalam, W., M. P. Devereux and G. Maffini (2009), "The Direct Incidence of Corporate Tax on Wages", Oxford University Centre for Business Taxation Working Paper 0917, University of Oxford
- Attanasio, Di Maro, Lechene and Phillips (2010), "The Impact of Food Price Rises on the Welfare of Poor Households in Mexico and Colombia", IFS Mimeo (forthcoming working paper).
- Banco de Mexico (2009), "Informe anual", ABRIL, 2010.
- Banks, J., R. Blundell and A. Lewbel (1997), "Quadratic Engel Curves and Consumer Demand", The Review of Economics and Statistics, Vol. 79, No. 4, pp. 527 - 539.
- Bell, K., M. Brewer and D. Phillips (2007), "Lone parents working under 16 hours per week ('mini-jobs')", The Joseph Rowntree Foundation
- Blundell, R. (2009), "Assessing the Temporary VAT Cut Policy in the UK", Fiscal Studies, Volume 30, Issue 1, pages 31-38, March 2009
- Blundell, R., L. Pistaferri and I. Preston (2007), "The transmission of income inequality into consumption inequality", Mimeo, University College London.
- Blundell, R. and I. Preston (1998), "Consumption Inequality and Income Uncertainty", The Quarterly Journal of Economics, Vol. 113, No. 2, pp. 603 - 640.
- Bosch, M. and W. F. Maloney (2010), "Comparative analysis of labor market dynamics using Markov processes: An application to informality", Labour Economics, Vol. 17, No. 4, pp. 621 - 631.
- Botero, J. C., S. Djankov, R. La Porta, F. Lopez de Silanes and A. Shleifer (2004), "The Regulation of Labor", The Quarterly Journal of Economics, Vol. 119, No. 4, pp. 1339 - 1382.

Brambila-Macias, J. (2008), "The Dynamics of Parallel Economies. Measuring the Informal Sector in Mexico", Mimeo, available at: <http://mpr.ub.uni-muenchen.de/8400/>.

Brewer, M., A. Duncan and A. Shephard (2007), "Did working families' tax credit work? The impact of in-work support on labour supply in Great Britain", *Labour Economics*, Vol. 13, No. 3, pp. 699 - 720

Brewer, M., R. May and D. Phillips (2010), "Taxes, Benefits and the National Minimum Wage", A Report for the Low Pay Commission, available at: http://www.lowpay.gov.uk/lowpay/research/pdf/FromLPC_Document_Feb.pdf.

Browne, J. and D. Phillips (2010), "Tax and benefit reforms under Labour", Institute for Fiscal Studies Election Briefing Note 1, London: Institute for Fiscal Studies, available at: <http://www.ifs.org.uk/bns/bn88.pdf>.

Carare, A. and S. Danninger, (2008), "Inflation smoothing and the modest effect of VAT in Germany", International Monetary Fund (IMF), Working Paper no. WP/08/175

Cebula, R. (1997), "An Empirical Analysis of the Impact of Government Tax and Auditing Policies on the Size of the Underground Economy: The Case of the United States, 1993 - 1994", *American Journal of Economics and Sociology*, Vol. 56, No. 2, pp. 173 - 185.

Centro de Estudios de las Finanzas Públicas (2009a) "Modelo de Microsimulación del Impuesto Sobre la Renta - Retención de Salarios - Nota Metodológica". CEFP, Palacio Legislativo de San Lázaro, agosto 2009.

Centro de Estudios de las Finanzas Públicas (2009b) "Simulador de la Incidencia del Impuesto al Valor Agregado. Nota Metodológica". CEFP, Palacio Legislativo de San Lázaro, septiembre 2009.

Centro de Estudios de las Finanzas Públicas (2009c) "Simulador de la Incidencia del Impuesto Especial sobre Producción y Servicios en Bienes y Servicios No Petroleros. Nota Metodológica". Notacefp / 053 / 2009. CEFP, Palacio Legislativo de San Lázaro, septiembre 2009.

Centro de Estudios de las Finanzas Públicas (2009d) "Análisis de la Incidencia de los Impuestos al Consumo en la Reforma Fiscal 2010: Iniciativa vs. Dictámenes". CEFP / 081 / 2009. CEFP, Palacio Legislativo de San Lázaro, noviembre 2009.

Centro de Estudios de las Finanzas Públicas (2009e) "Incidencia de la Miscelánea fiscal 2010 aprobada por el H. Congreso de la Unión" CEFP / 082 / 2009. CEFP, Palacio Legislativo de San Lázaro, noviembre 2009.

Das, M., and A. van Soest (2001), "Family labour supply and proposed tax reforms in the Netherlands", *De Economist*, Vol. 149, pp. 191 - 218

Emmerson, C, H. Reed and A. Shephard (2004), "An Assessment of PENSIM2", Institute for Fiscal Studies Working Paper 04/21, London: Institute for Fiscal Studies, available at: <http://www.ifs.org.uk/wps/wp0421.pdf>.

Fullerton, D, and G. Metcalf (2002), "Tax Incidence", Chapter 26, *Handbook of Public Economics*, Volume 3, Eds. A.J. Auerbach and M. Feldstein, Amsterdam: Elsevier.

- Galiani, S. and F. Weinschelbaum (2007), "Modelling Informality Formally: Households and Firms", CEDLAS, Universidad Nacional de La Plata, Working Paper No. 0047
- Gong, X. and A. van Soest (2002), "Family Structure and Female Labor Supply in Mexico City", *Journal of Human Resources*, Vol. 37, No. 1, pp. 163 – 191.
- Gruber, J. (1997), "The incidence of payroll taxation: evidence from Chile", *Journal of Labor Economics*, Vol. 15, No. 3, pp. S72 – 101.
- Levy, S. (2008), "Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico", Washington, DC: Brookings Institution Press.
- Lopez-Calva, L.F., Á. Méndez Martínez, E.G. Rascón Ramirez, L. Rodríguez-Chamussy, and M. Székely Pardo (2008), "El ingreso de los hogares en el mapa de México". *El Trimestre Económico*, 2008, vol. LXXV (4), issue 300, pages 843-896
- Lopez-Calva, L.F., Á. Méndez Martínez, E.G. Rascón Ramirez, L. Rodríguez-Chamussy, and M. Székely Pardo (2007), "Poniendo a la pobreza de ingresos y a la desigualdad en el mapa de México". *Economía Mexicana NUEVA EPOCA*, 2007, vol. XVI, issue 2, pages 239-303
- Maloney, W. F. (1999), "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico", *The World Bank Economic Review*, Vol. 13, No. 2, pp. 275 – 302
- Maloney, W. F. (2004), "Informality Revisited", *World Development*, Vol. 32, No. 7, pp. 1159 - 1178
- Martini, A. and U. Trivellato (2003), "The role of survey data in microsimulation models for social policy analysis", *Labour*, Vol. 11, pp. 83 – 112.
- McKay, S. F. (2003), "Dynamic Microsimulation at the U.S. Social Security Administration", Conference Paper, International Microsimulation Conference on Population, Ageing and Health 2003.
- Meghir, C., R. Narita and J. Robin (2008), "Formal work, informal work and wages in Brazil: an equilibrium model with search", Mimeo, University College London
- Meghir, C. and D. Phillips (2010), "Labour supply and Taxes", Chapter 3, *Dimensions of Tax Design*, Eds: J. Mirrlees et al, Oxford: Oxford University Press.
- Mirrlees, J., S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba (2010), "Tax by Design: The Mirrlees Review", Oxford: Oxford University Press for the Institute for Fiscal Studies
- Nigenda, G. (2005), "El seguro popular de salud en México Desarrollo y retos para el future". *Nota Técnica de Salud No. 2/2005*, April 2005, Banco Interamericano de Desarrollo
- Phillips, D. (2009), "Food price rises in Mexico: What is the value of detailed demand system analysis?", Institute for Fiscal Studies, Mimeo
- Plumb, M. (2001), "An Integrated Microsimulation and Applied General Equilibrium Approach to Modelling Fiscal Reform", Conference Paper, Econometric Society Australasian Meeting 2001.

Samaniego, R. et al (2006), "Medición de la Evasión Fiscal en México", CAEPP, Instituto Tecnológico Autónomo de México

Schaefer, T. and A. Peichl (2006), "Documentation FiFoSiM: integrated tax benefit microsimulation and CGE model", FiFo-CPE Discussion Paper 06-10, University of Cologne and Cologne Centre for Public Economics

Schneider, F. and D. H. Enste (2000), "Shadow Economies: Size, Causes and Consequences", *Journal of Economic Literature*, Vol. 38, No. 1, pp. 77 – 114.

Schneider, F. (2002), "Size and Measurement of the Informal Economy in 110 Countries Around the World", World Bank Research Project, Washington D.C.: World Bank

Scott, J. (2006), "Seguro Popular: Incidence Analysis". September 2006, CIDE, Working Paper No. 367.

Secretaría de Hacienda y Crédito Público (SHCP) (2010), "Distribución del pago de impuestos y recepción del gasto público por deciles de hogares y personas"

Ulyssea, G. (2010), "Regulation of entry, labor market institutions and the informal sector", *Journal of Development Economics*, Vol. 91, No. 1, pp. 87 - 99

van Soest, A. (1995), "Structural Models of Family Labor Supply: A Discrete Choice Approach", *The Journal of Human Resources*, Vol. 30, No. 1, pp. 63 – 88

Vuletin, G. (2008), "Measuring the Informal Economy in Latin America and the Caribbean", IMF Working Paper, No. 08/102

Appendix A. A description of the Mexican tax system

In this section we detail those parts of the Mexican tax system that will be analysed and simulated in this project. For a description of the corporate tax system please see X.²¹ A description of the welfare system in Mexico can be found in Y. This section draws on the exposition of the tax system found in Absalón and Urzúa (2009a).

A1. The Value Added Tax (VAT / IVA)

The rate of VAT applicable in Mexico prior to the reforms to be considered in this project was 15%. In those towns bordering the US, a reduced rate of 10% was applied. Unprepared foods, medicines and contraceptives are zero-rated (i.e. no VAT is charged on the sale of these goods and sellers *can* reclaim any VAT paid on inputs), whilst public transport, educational goods and services, books, magazines, housing, and medical services are exempt (i.e. no VAT is charged on the sale of these goods but sellers *cannot* reclaim the VAT paid on inputs). Extensive zero-rating and exemptions, together with a large informal economy in which VAT (amongst other taxes) is evaded mean that the Mexican VAT system raises only 33% of the revenue that it would if the standard rate of VAT was applied to all final consumption. This is by far the lowest rate in the OECD, of which Mexico is a member, but is comparable to other middle-income Latin American countries.

Table A1 shows monthly expenditure as recorded by ENIGH (2008) by VAT category and expenditure group. It shows that spending on food VAT exempt or zero-rated goods totals 46.7% of total expenditure. A further 4.1% of spending consists of VATable goods purchased from retailers that we consider to be 'informal', with 49.2% of spending on VATable goods purchased at retailers we consider formal. This would imply that approximately 8% of VAT was evaded, a considerably lower figure than estimated by the government (see section 2.4), suggesting that a part of expenditure from 'formal' retailers is also subject to tax evasion.

Table A.1 Household Expenditure by VAT and formality classification

	Expenditure	% of total expenditure
VAT exempted		
Other non-food goods and services	14600.96	7.29
Health services	2065.8	1.03
Education	12703.66	6.35
Lottery	135.03	0.07
Public transport	9910.26	4.95
Transfers	3476.48	1.74

²¹ Please note that the new *Impuesto Empresarial a Tasa Unico* (IETU, or 'Single Rate Business Tax') will not be included in the indirect tax system for the purposes of this project or the MEXTAX simulator model, at least at this stage. This is because, although the tax itself operates as, in effect, a second Value Added Tax, presently it is complementary to the standard corporate income tax such that the higher of the two taxes is paid. Hence, it is not clear what the link between expenditure and the amount of IETU paid by consumers (through increased prices) is, unlike for standard VAT, for which an assumption of full pass-through to consumer prices is common. Furthermore, whilst the rate of IETU has been increased in 2010 (from 17.0% to 17.5%), this was a pre-planned change and not part of the 2009 congressional debates.

VAT zero rate		
Food	45185.51	22.57
Other non-food goods and services	3357.81	1.68
Health and Medicines	2010.17	1.00
VAT taxed		
Food	19275.14	9.63
Formal	16261.78	8.12
Informal	3013.36	1.51
Other non-food goods and services	65922.11	32.93
Formal	60656.43	30.30
Informal	5265.68	2.63
Health-related goods and services	1982.42	0.99
Petrol	9299.92	4.65
Telecoms	9302.89	4.65
VAT taxed and duties		
Alcohol	116.86	0.06
Under 14° alcohol	17.44	0.01
14°-20° alcohol	14.81	0.01
Over 20° alcohol	84.61	0.04
Beer	377.08	0.19
Tobacco	470.71	0.24
Total	200192.84	100

Notes: Figures are calculated using sampling weights (variable 'factor' in ENIGH 2008). Informal expenditure is defined as those goods and services purchased from informal vendors such as street markets and stalls (ENIGH variable lug_com equal to 1, 2, or 3). Expenditure data is at the household level, and includes expenditure from the ENIGH files 'gastos', 'gastodiario', and 'gastoeduca'. The classification of goods and services into the different taxation categories corresponds to the VAT and Duties systems valid in 2008; it is based on CEFP (2009b, 2009c). Monthly expenditure figures are shown. Source: ENIGH 2008

A2. Excise Duties (*IEPS*)

Mexico operates a fairly complex system of duties where some are ad-valorem, whilst others consist of specific duties.

For alcohol, tax is levied as a proportion of the pre-tax sale price of the alcohol. Under the pre-reform tax system, if the percentage of alcohol by volume is less than 14%, this rate is 25%, where the percentage of alcohol is between 14% but less than 20%, the rate is 30%. Otherwise the rate is 50%. Hence, most beers and wines are taxed at a rate of 25%, and most spirits at 50%. A minimum fee of 3 pesos per litre is applied to 'avoid the use of (cheaper) disposable containers', and in addition, the amount to be paid can be reduced by 1.26 pesos if a reusable container is used.

For tobacco, the duty rate is 160% in general, with a reduction to 30.4% for hand-made cigars.

For gasoline and diesel, SAT publishes the IEPS rates monthly and for each selling agency. To calculate these rates, SAT uses the international oil price, the price set by PEMEX, and the price paid by the consumers (which is set by the government). Sometimes the rates can be negative, for example as a consequence of an increase in the international price of oil as in 2008. When rates are negatives, according to the article 2-A in the IEPS Law 2008, PEMEX and its subsidiaries will be able to deduct the resulting amount from their IEPS or VAT debit; hence insulating the consumer from sharp rises in oil prices.

For diesel the flat-rate amount was 29.88 cents by litre in 2008. For petrol (magna) the flat-rate amount was 36.00 cents per litre and for petrol (premium), 43.92 cents per litre.

In our analysis, we shall use the % rates and ignore the flat rate minima and amounts, except for petrol and diesel.

Table A.2 Excise Duties (IEPS) Rates in 2008

Good	% of price	Flat-rate amount
Beer (and alcohol less than 14%)	25%	0
Alcohol 14 – 20%	30%	0
Alcohol 20% +	50%	Minima apply
Cigarettes	160%	0
Other Tobacco	160%	0
Lottery Tickets	20%	0
PEMEX Diesel*	-59.30%	29.88c/l
Diesel low sulphur*	-52.95%	29.88c/l
Petrol PEMEX Magna*	-38.69%	36.00c/l
Petrol Premium*	-29.19%	43.92c/l
Telecoms	0%	0

Notes: *Corresponding to August 2008 and selling agency Azcapotzalco²²

A3. Social Security Contributions

Mexico operates separate social security systems for the private sector (the *Instituto Mexicano del Seguro Social (IMSS)*) and the public sector (the *Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE)*). The pensions component of the IMSS was converted to individual-accounts with compulsory contributions for contributions made after 1997, with similar changes enacted for the ISSSTE for contributions made after 2007. For the purposes of this project, however, these contributions are treated as traditional payroll taxes. Whilst we make note of government contributions to the social security funds, these are not modelled as part of this project as the micro-simulation tool that will be developed looks at contemporaneous income as opposed to lifetime resources (which would include contributions to pensions from the government's general tax revenues). Exemptions for certain forms of income are similar to those for income tax (see section A4) and for the purposes of this project, we model exemptions as if they were the same.

²² See SAT website:
http://www.sat.gob.mx/sitio_internet/asistencia_contribuyente/informacion_frecuente/ieps/41_12273.html

As well as the contributions set out below, the employer contributes an amount equal to 5% of the wage to funds known as *INFONAVIT* (private sector) or *FOVISSSTE* (public sector) that can be used towards the purchase of a house (or if unused become part of the pension fund).

A3.1. A description of *IMSS*

The *IMSS* provides social security benefits to workers in the private sector and covered 48 million people (including contributors and their dependents) as of 2008. The contributions as currently set out in law are found in table A.3. Self-employed persons can contribute voluntarily at a flat-rate of one minimum wage salary. As detailed in the income tax section, additional voluntary contributions to the pension fund can be made tax free (subject to a limit).

**Table A.3 Mandatory social security contributions (% of salary unless otherwise stated)
(Rates as of 2008)**

Contribution Type	Employer	Employee	Federal Government
Retirement Pension	2.00%	0.00%	5.50% Federal District Minimum Wage
Unemployment in advanced age	3.15%	1.125%	0.00%
Disability and life insurance	1.75%	0.625%	7.143% of the employer's contribution
Sickness and maternity leave	0.70%	0.25%	0.05%
Sickness and maternity expenses	6.00% ^a	2.0% ^a	-
Benefits-in-kind	1.05%	0.375%	0.075%
Child-care and social services	1.00%	-	-
Job-risk contribution	0.5% - 15%	-	-

Notes: a: These rates apply to earnings in excess of three times the minimum wage only. In addition, employers pay a flat rate equal to 13.9% of a minimum wage salary irrespective of the level of earnings.

Sources: Urzua and Absalon (2009a).

A3.2. A description of *ISSSTE*

The *ISSSTE* provides social security benefits to workers in the public sector and covered 11 million people (including contributors and their dependents) as of 2008. The contributions as currently set out in law are found in table A.4

Table A.4 Mandatory social security contributions (% of salary unless otherwise stated)
(Rates as of 2008)

Contribution Type	Employer	Employee	Federal Government
Retirement Pension	2.00%	0.00%	-
Unemployment in advanced age	3.175+%	5.075%	-
Disability and life insurance	0.625%	0.625%	-
Sickness and maternity leave	7.375%	2.75%	-
Child-care and social services	0.5%	0.5%	-
Job-risk contribution	0.75%	-	-

Sources: Urzua and Absalon (2009a).

A4. Personal Income Tax

The personal income tax (*ISR*) is payable by residents of Mexico (who pay tax on their world-wide income) and non-residents (who pay tax on income attributable to Mexico). Taxes due on wage earnings are deducted automatically by the employer through withholding, and provisional monthly payments must be made for tax accruing on unearned income. The final bill is calculated after the end of the fiscal year, with further payments levied or refunds given as necessary.

Income from wages, pensions, benefits and financial capital forms the tax base for the *ISR*; self-employment income and property income is treated differently, and is subject to the *IETU*, Mexico's single rate business tax. The tax simulators designed by CIEP and Absalón and Urzúa consider only employment income. In the first instance, our model shall be restricted to this set of taxes, but it is planned to extent the model to cover unearned income if time and resources permit.

A fairly complex system of exemptions operates:

- Half of the earnings from overtime work are exempt (100% if the worker receives at most the minimum wage) provided that the amount of overtime worked does not exceed 3 hours per day for at most three days per week, and that the weekly exemption totals less than five minimum salaries for the week.
- Social security benefits from *IMSS* and *ISSSTE* (such as pensions and disability pensions) are exempt up to nine times the monthly minimum wage salary. Other social security benefits are exempt up to one times the monthly minimum wage salary.
- Employer-provided retirement pensioners and associated benefits are exempt provided they do not exceed ninety times the monthly minimum wage salary per year of service.
- Profit-sharing and vacation allowances are exempt for at most 15 times the daily minimum salary.

- Annual bonuses are exempt up to an amount equal to a monthly minimum wage salary
- All interest paid on savings deposits is exempt.

We will model these exemptions as fully as possible, using the work of CIEP and Absalón and Urzúa as a guide.

There are also a set of deductions for certain expenditures:

- Voluntary contributions to individual pension accounts (AFORES) up to an amount equal to 10% of income.
- Charitable donations up to an amount equal to 7% of income.
- Funeral expenses up to an amount equal to the annual minimum wage.
- School bus transport
- Medical insurance
- Medical services

The tax simulator produced by CIEP does not account for these deductions. Initially, our simulator will abstract from these deductions because they increase the complexity of calculating gross income significantly (see appendix C). However, if time permits these deductions will be included, and modelled, assigning expenditure on these goods to the individual with the highest taxable labor income.

Tax is applied from the first peso of taxable income (subject to the above exemptions and deductions), using a progressive rate schedule (see table A.5). A refundable tax-credit, the *Subsidio para el Empleo* (subsidy for employment) means that those with low earnings could be liable for a negative amount of income tax. Table A.6 shows the structure of this tax credit, which is withdrawn in steps as income increases, as opposed to being withdrawn via a smooth rate schedule.

Table A.5 Personal Income Tax Schedule (2008)

Income Tax Rate	Rate Threshold (pesos per annum)	Rate Limit (pesos per annum)
1.92%	0.01	5952.84
6.40%	5952.85	50524.92
10.88%	50524.93	88793.04
16.00%	88793.05	103218.00
17.92%	103218.01	123580.20
19.94%	123580.21	249243.48
21.95%	249243.49	392841.96
28.00%	392841.97	-

Sources: Urzua and Absalon (2009a).

Table A.6 *Subsidio para el Empleo Schedule (2008)*

Wage Threshold	Wage Limit	Subsidy amount
0.01	21227.52	4884.24
21227.53	31840.58	4881.96
31840.57	41674.08	4879.44
41674.09	50524.92	4713.24
50524.93	53353.80	4589.52
53353.81	56606.16	4250.76
56606.17	64025.04	3898.44
64025.05	74696.04	3232.56
74696.05	85366.80	3042.48
85366.81	88587.96	2611.32

Sources: Urzua and Absalon (2009a).

A5. Budget constraints and the distribution of tax payers in ENIGH

Figures A1 and A3 show the combined marginal and average income tax and social security contribution rates for those covered by IMSS and ISSSTE, and figures A2 and A4 show the associated budget constraints (i.e. net income for a given taxable income). These figures consider employee social security contributions only.²³

The first thing to note is that for both IMSS and ISSSTE, whilst at low levels of income, marginal rates increase with income, this is not the case at high levels as the social security contributions limit is reached. This occurs at \$479,883.75 (Mex) for IMSS and \$191953.50 (Mex) for ISSSTE. At these points combined marginal rates fall.

Second, the withdrawal of tax credits in discrete amounts rather than via a smooth schedule results in extremely high combined marginal rates at some earnings levels. This is most notable when earnings reach \$88,587.96 (Mex) when the tax credits are finally withdrawn and one peso of extra income leads to the loss of \$2,611.32 (Mex). The average tax rate jumps from 13.6% to 16.5% at this point.

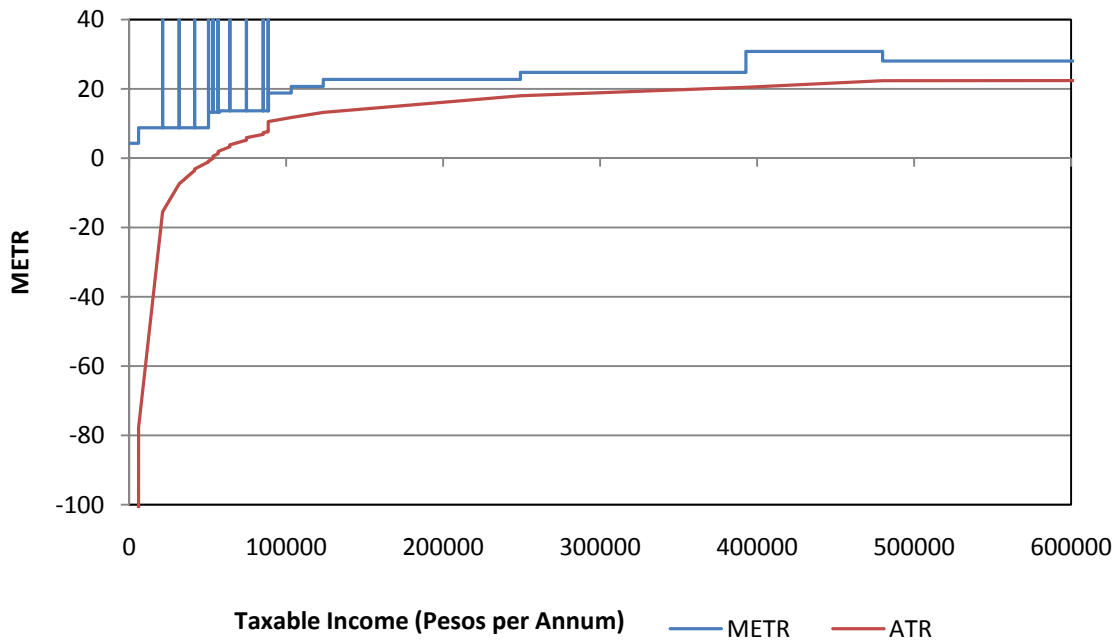
The average tax-rate is lower than the marginal tax rate at all levels of earnings for those subject to IMSS, but not for those subject to ISSSTE. This is because the ISSSTE contributions limit is reached when the marginal income tax rate is 19.94% (as opposed to 28%) and because the rate of ISSSTE contributions is over 3 times that for IMSS. Higher contribution rates for ISSSTE mean that most workers covered by this scheme have higher marginal tax rates (although those earning between \$191953.50 (Mex) and \$479,883.75 (Mex) have a lower marginal combined tax rate as contributions are still required by IMSS but not by ISSSTE at these income levels). Average combined tax rates are higher under ISSSTE for all earnings, however.

The average tax rate is negative for low earnings due to the tax credit. This is true for earnings up to about \$54,000 (Mex) for IMSS and \$35,000 (Mex) for ISSSTE.

²³ Including employer contributions is possible, but on the X-axis rather than taxable income one would have employer cost (under the assumption that no exemptions apply)

Table A7 shows the distribution of tax-payers across tax bands both before and after increasing labor income by the 'Altimir factor' required for ENIGH figures to match National Accounts. Under the former, the most common marginal rate for earned income is 6.4%, and under the latter 10.88%.

Figure A1: Income Tax and IMSS Marginal Effective Tax Rate (METR)



Source: Urzua and Absalon (2009a) and authors' calculations. Also source for A2 to A4.

Figure A2: Net income under the IMSS system

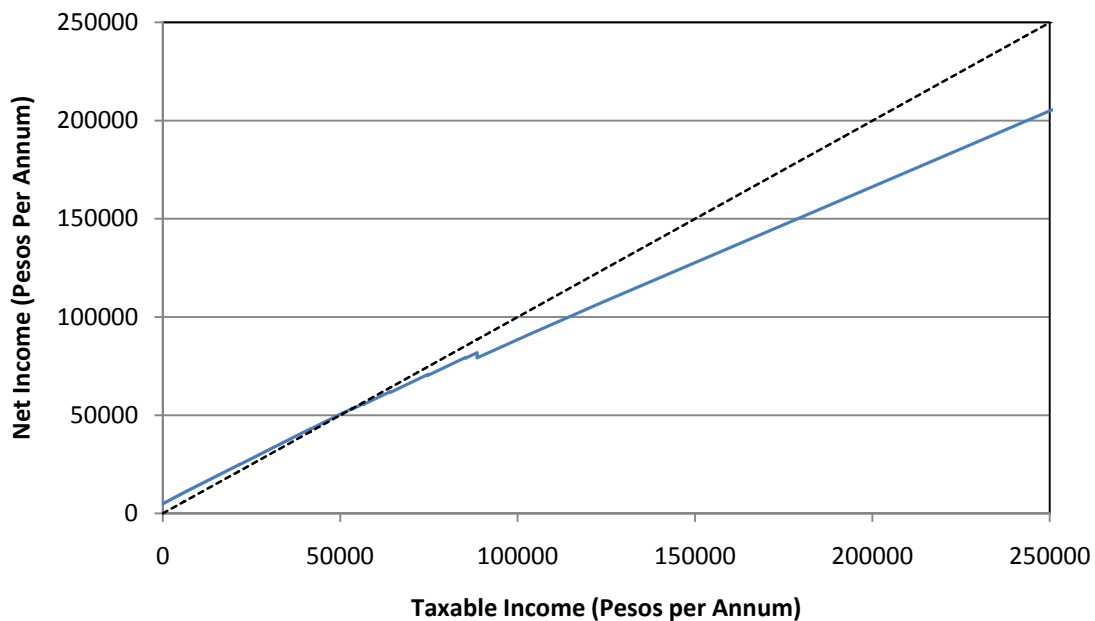


Figure A3: Income Tax and ISSSTE Marginal Effective Tax Rate (METR)

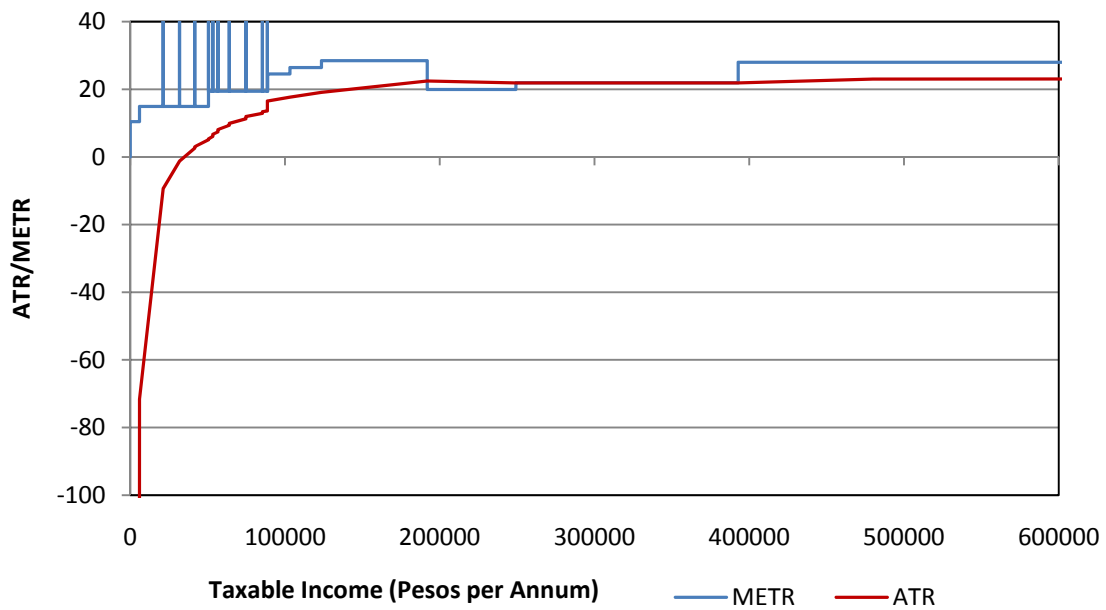


Figure A4: Net income under the ISSSTE system

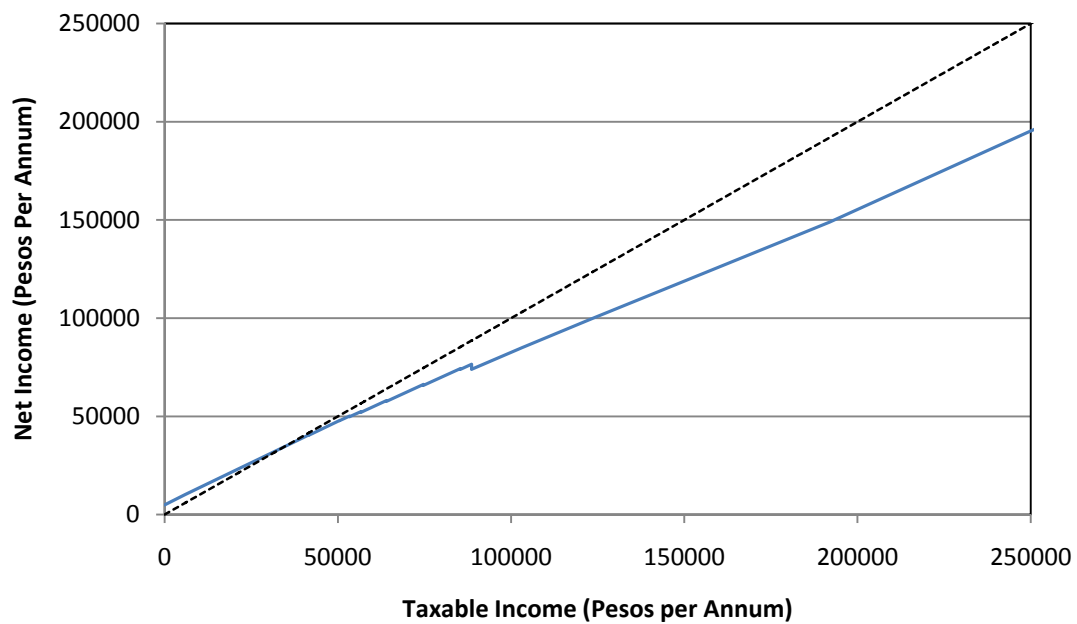


Table A.7 Distribution of Taxpayers across the Taxable Income Bands (Employment Income)

Income Tax Rate	Proportion of Taxpayers (Raw ENIGH data)	Proportion of Taxpayers (Adjusted ENIGH data)
1.92%	0.0%	0.0%
6.40%	38.1%	19.7%
10.88%	29.9%	28.2%
16.00%	5.8%	8.8%
17.92%	6.5%	10.4%
19.94%	12.8%	19.5%
21.95%	5.2%	9.0%
28.00%	1.9%	4.5%

Notes: 'Adjusted ENIGH data' is where all gross incomes are adjusted by a constant factor. No taxpayers are recorded as paying the 1.92% tax-rate at the margin because we assume that formal workers are paid at least a minimum wage salary in line with the analysis of CIEP/CEPF.

Source: ENIGH (2008) and authors' calculations.

Appendix B. The planned and implemented tax reforms

This section details the 2010 income tax and indirect tax reforms as initially proposed by the Federal Government and as finally approved by both houses of the Mexican congress and implemented. It provides not only a description of the policy but also a brief analysis of its likely impact on behavior and across the income distribution.

B1. The Initial Proposals for the 2010 Fiscal Reform Package

The Federal Government's initial proposals consisted of:

- The introduction of a new general expenditure tax applying at a rate of 2.0% on all goods and services, including those currently exempt and zero-rated under existing value-added tax legislation.
- An increase in the tax on cash deposits from 2.0% of the balance to 3.0% of the balance.
- An increase in the top rate of income tax (both personal and corporate) from 28.0% to 30.0% in 2010, 2011 and 2012, with a phased reduction to 28.0% in 2014.
- Increases in the second and third highest rates of income tax from 21.95% to 23.52%, and 19.94% to 21.36%, respectively. A reduction in the threshold of the 16% tax rate.
- An increase in the minimum charge per litre of 3 pesos for all drinks with an alcohol content greater than 20%.
- An increase in the rate of tax on beer from 25% to 28%
- Introduction of a tax of 0.04 pesos per cigarette (0.75 grams of snuff), to be increased to 0.10 pesos by 2014.
- An increase in the tax on lottery games by 10% to 30%.
- Introduction of a 4% tax on telecommunications services.

The Federal Government planned that these additional revenues would be used to fund increases in the coverage of various social programmes, notably Oportunidades, and to strengthen the public finances by diversifying revenue away from the oil sector. The description of the distributional and behavioral effects of these changes is given in Section B3, allowing a more natural comparison with the changes implemented following debates in the Mexican congress.

B2. The Approved 2010 Fiscal Reform Package

Following debates in the upper and lower chambers of the Mexican congress, a final set of reforms was approved which includes the following reforms:

- An increase in the standard rate of VAT/*IVA* to 16.0% (11.0% in border areas).
- An increase in the tax on cash deposits from 2.0% of the balance to 3.0% of the balance.
- An increase in the top rate of income tax (both personal and corporate) from 28.0% to 30.0% in 2010, 2011 and 2012, with a phased reduction to 28.0% in 2014.
- Increases in the second and third highest rates of income tax from 21.95% to 23.52%, and 19.94% to 21.36%, respectively.
- An increase in the tax rate on drinks containing more than 20% alcohol, by volume, from 50% to 53%.
- A temporary increase in the rate of tax on beer from 25% to 26.5%

- An increase in the tax on lottery games by 10% to 30%.
- Introduction of a 3% tax on telecommunications services.

This represents a smaller overall tax rise than that initially proposed, with the Federal Government estimating additional fiscal revenues as a proportion of GDP of around 1%²⁴.

B3. The distributional and behavioral impact of reforms

Broadly-speaking, whilst raising less revenue, the approved package is more progressive than the initial proposals (CEPF have conducted quantitative analysis showing this, see CEPF (2009)). This is mainly because the relatively narrow base of the Mexican VAT makes up a larger proportion of the expenditure of richer households than for poorer households (food, a necessity, is the largest component of those goods not subject to the standard rate), so an increase in VAT has a greater proportional effect on the rich, whereas the proportional effect would be the same for an increase in a universal consumption tax.

Increases in the top rates of income tax will affect a minority of (high income) Mexicans. For instance, around one third of formal workers pay the top three rate of income tax on their labor income according to (adjusted) ENIGH figures (the figure is around one fifth according to adjusted ENIGH data). Evidence from developed economies suggests that higher earners are particularly responsive to changes in taxation because of the high degree of autonomy they have over their work effort and their remuneration. As well as reductions in effort, higher taxes will act to encourage tax evasion and avoidance activities. As detailed in section 3.4, this is an issue that we think requires significantly more work than can be undertaken in this project, but that we plan to begin to address using a series of sensitivity and robustness analyses.

The increases in the income tax are not permanent and operate, in full, for 3 years only, with the top rate of income tax reverting to its 2009 level in 2014. This temporary nature may affect the magnitude of the impact, but the direction of this effect is unclear. For instance, the decisions to move from the formal to informal sector may be muted if there are frictions involved, but responses of total labor supply or income realisation in the short-run may be larger as inter-temporal substitution takes place. The extent to which the pledge to reduce the tax back to its 2009 levels in the future is understood and believed is also not known. For these reasons we will not formally consider the inter-temporal aspects of labor supply response in this project.

Increases in the consumption taxes and duties will affect more people. Virtually all Mexicans would be affected by increases in VAT (except those purchasing solely from the informal economy). According to ENIGH, spending on beer accounts for 0.19% of total expenditure (with 4% of households purchasing at least some beer), other alcohol for 0.06% (1% of households), tobacco 0.24%, lottery tickets for 0.07% (7% of households), and telecommunications for 4.65% (71% of households). The expenditure and number of households purchasing alcohol, tobacco and lottery tickets are very significantly under-recorded in ENIGH, however and adjustments will need to be made to account for this.

An increase in the existing rate of VAT as opposed to the introduction of a general consumption tax is likely to cause bigger shifts in relative prices, and hence a larger response amongst consumers' expenditure patterns to changes in these prices. However, offsetting this is the fact

²⁴ See Bank of Mexico, Annual Report 2009, page 40.

that the fact that as the overall increase in taxes is lower, the 'income effect' of the tax reforms will be smaller under the adopted reforms compared to the proposed reforms. Without knowledge of the sensitivity of consumer demands to income and prices, that makes it difficult to assess whether expenditure patterns will be affected more by the adopted reforms than those initially proposed. We plan to address these issues in this project. Any increase in consumption taxes will also encourage a shift from the formal to the informal sector of the economy (for instance from retail stores to travelling markets). It is also important to recognise that expenditure taxes reduce work incentives in broadly the same way as direct taxes. By raising the overall effective rate of taxation, the proposals will also act to discourage work. These issues shall not be addressed in this paper.

APPENDIX C: DESCRIPTION OF DATA CREATION

The main data used in the analysis will be from ENIGH 2008. Using this survey we will construct the main input files to be used in the MEXTAX simulator, combining information from different files contained in ENIGH 2008 raw data. Here we describe each of these files in more detail and provide details on how we construct the variables that are likely to be contained in the input files; how we calculate gross labor income; and how we will create alternative input files to conduct our sensitivity analysis.

Monthly expenditure figures are calculated dividing quarterly expenditure figures (variable `gas_tri` in files `gastos`, `gastodiario` and `gastoeduca`; variable `apo_tri` in file `nomonetario`) provided in ENIGH by 3. As in previous existing work, we use the variable `lug_com` to classify expenditure into formal and informal, in the consumption file described below. Informal expenditure comprises purchases from informal vendors such as street markets as defined by `lug_com` equals 1, 2 or 3. The classification of goods and services into the different taxation categories corresponds to the VAT and Duties systems valid in 2008; it is based in CEFP (2009b). The consumption file contains only monetary expenditure.

Monthly income figures are calculated taking into account the period in which the survey was applied to each particular household. For each income source, the quarterly figure is divided by one of the following numbers: 2.99178 if `decena` equals 1; 3.02465 if `decena` equals 2; 3.02465 if `decena` equals 3; 3.02465 if `decena` equals 4; 3.00821 if `decena` equals 5; 3.00821 if `decena` equals 6; 3.00821 if `decena` equals 7; 3.02465 if `decena` equals 8; 3.02465 if `decena` equals 9.

The baseline modelling will use raw ENIGH data on incomes and expenditures but the sensitivity analysis will require adjustment for under-reporting of income and expenditure by grossing income up by constant or varying factors. Household members' report their net income, after paying their personal income taxes and making their social security contributions, in the ENIGH survey. We calculate gross labor income figures using the CIEP simulator that recovers the pre-tax income, or gross labor income, for each individual. The model is based on the Mexican Income Tax Law and Social Security Law prevailing in 2008. To get gross labor income for each household, individuals' incomes are grouped into these categories: wages and salaries; overtime; end-of-the-year bonus; incentives, rewards and prizes; holiday bonuses and allowances in cash; profit sharing from secondary subordinated work and end-of-the-year bonus; income from pensions; and other labor income. For each individual we will calculate the total net income that was taxed. Initially we will use the tax rates and social security contributions described in tables A.3 to A.6 to calculate gross income using total taxable net income. We will then calculate an average effective tax rate (the ration between total tax paid and total gross income) and apply this to each income source to get gross income for each of these. Once we understand the order of taxation we will calculate gross income separately for each source using the marginal rates applying for that component of income based on the order in which it is taxed.

As mentioned before, we use the following definition of formal worker as a baseline to classify income into formal and informal in the adult file: formal workers are individuals with a positive amount of (net) income through labor and receiving any of the following social benefits (as provided by the ENIGH file `pobla08`): IMSS (`inst_1=1`), ISSSTE (`inst_2=2`), state ISSSTE (`inst_3=3`) or PEMEX (`inst_4=4`), through work (variable `inscr_1=1`). Contributions to social

security differ for workers covered by the private sector health service (IMSS) or by the public sector (ISSSTE, state ISSSTE or PEMEX) as described in Appendix A. When an individual appears to be covered by IMSS and any of the public sector health systems, we use information from the ENIGH file trabajos about whether the worker's main job is in the public or private sector (variables clas_emp and numtrab=1).

We initially assume that all workers comply with all her tax and social security obligations, that taxable income comes from principal and/or secondary employment, and that the tax impact falls entirely on the worker. This will mean that the baseline input files will assume that labor income is either all formal (and therefore subject to tax and social security contributions) or all informal (and therefore not subject to tax and social security contributions) for each worker. We will also assume that all formal workers receive at least a minimum daily wage.

Household file

Variables	Definition	ENIGH 2008 file	Variable description
folioviv	Residence identifier		
foliohog	Household identifier		
residentes	Number of household members	hogares	
menores	Number of household members under 12 years old	hogares	
mayores	Number of adult household members		residentes-menores
eqsc_5030	Equivalence scale		$1 + (\text{mayores}-1) \cdot (0.5) + \text{menores} \cdot (0.3)$
eqsc_7045	Equivalence scale		$1 + (\text{mayores}-1) \cdot (0.7) + \text{menores} \cdot (0.45)$
eqsc_8055	Equivalence scale		$1 + (\text{mayores}-1) \cdot (0.8) + \text{menores} \cdot (0.55)$
eqsc_9060	Equivalence scale		$1 + (\text{mayores}-1) \cdot (0.9) + \text{menores} \cdot (0.6)$
factor	Sampling weights	hogares	
estrato	Stratum	hogares	
ubica_geo	Area Identifier	hogares	
totexp_month	Total expenditure (monthly)	gastos, gastodiaros, gastoeduca, nomonatarario, hogares	Sum of all expenditure categories + imputed rent
totmonexp_month	Total monetary expenditure (monthly)	gastos, gastodiaros, gastoeduca	Sum of all expenditure categories
totnmonexp_month	Total non monetary expenditure (monthly)	nomonatarario, hogares	Sum of all non monetary expenditure + imputed rent
net_totmoninc_month	Total net monetary income (monthly)	ingresos	Sum of all income sources in the household
net_totinc_month	Total net income (monthly)	ingresos , nomonatarario	net_totmoninc_month +totnmon_exp_month

Consumption file

Variables	Formal/ Informal	Tax classification 2008	Definition	ENIGH 2008 file	Variable description
folioviv			Residence identifier		
foliohog			Household identifier		
exp1	Formal	VAT exempted	Health services	Gastos	Clave=J001, J004, J005, J007, J008, J013, J016, J017, J018, J036, J039, J062, J072
exp2	Formal	VAT exempted	Education services	Gastoeduca	Clave=E001-E007
exp3	Formal	VAT exempted	Lottery	Gastos	Clave=E029
exp4	Formal	VAT exempted	Public transport	Gastodiario	Clave=B001-B007
exp5	Formal	VAT exempted	Transfers	Gastos	Clave=N011, N012, N013, N014, N015, N016
exp6	Formal	VAT exempted	Other non-food goods and services	Gastos	Clave= E008 E013 E015 E017, E020, E021, E022, E026, G001-G006, G011- G019, H120-H136, M001, N006, N007, T902, T905
exp7	Formal	VAT zero rate	food	Gastodiario	
exp8	Formal	VAT zero rate	Other non-food goods and services	Gastos	Clave= E014, G007, L029, T910
exp9	Formal	VAT zero rate	Health-related goods and services	Gastos	Clave=J009, J010, J014, J020-J035, J037, J038, J042, J044-J059, J061, J063, J064
exp10	Formal	VAT taxed	Food	Gastodiario	XXXX
exp11	Formal	VAT taxed	Other non-food goods and services	Gastos	Clave= C001-C024, D001-D026, E009- E012, E016, E018, E019, E023-E025, E027, E028, E030- E033, F007, F013- F017, G008-G010, G020-G022, H001- H119, I001-I026,

					K001-K044, L001-L028, M002-M018, N001-N005, N008-N010, T901, T903, T904, T906-T909, T911-T915
exp12	Formal	VAT taxed	Health-related goods and services	Gastos	Clave= J002, J003, J006, J011, J012, J015, J019, J040, J041, J043, J060, J065-J071
exp13	Formal	VAT taxed	Petrol		Clave= F010, F011, F012
exp14	Formal	VAT taxed	Telecoms		Clave= F001 , F002, F003, F004, F005, F006, F008, F009
exp15	Formal	VAT and duties	under 14° alcohol	gastodiario	Clave=A228, A231 ,A234, A232, A238
exp16	Formal	VAT and duties	14°-20° alcohol	Gastodiario	Clave= A226, A237
exp17	Formal	VAT and duties	over 20° alcohol	Gastodiario	Clave = A223, A225, A227, A229, A230, A233, A235, A236
exp18	Formal	VAT and duties	beer	Gastodiario	Clave= A224
exp19	Formal	VAT and duties	tobacco	Gastodiario	Clave= A239, A240
exp20	Informal	VAT exempted	Health services	Gastos	Clave=J001, J004, J005, J007, J008, J013, J016, J017, J018, J036, J039, J062, J072
exp21	Informal	VAT exempted	Education services	Gastoeduca	XXXX
exp22	Informal	VAT exempted	Lottery	Gastos	Clave=E029
exp23	Informal	VAT exempted	Public transport	Gastodiario	Clave=B001-B007
exp24	Informal	VAT exempted	Transfers	Gastos	Clave=N011, N012, N013, N014, N015, N016
exp25	Informal	VAT exempted	Other non-food goods and services	Gastos	Clave= E008 E013 E015 E017, E020, E021, E022, E026, G001-G006, G011-G019, H120-H136, M001, N006, N007,

					T902, T905
exp26	Informal	VAT zero rate	food	Gastodiario	
exp27	Informal	VAT zero rate	Other non-food goods and services	Gastos	Clave= E014, G007, L029, T910
exp28	Informal	VAT zero rate	Health-related goods and services		Clave=J009, J010, J014, J020-J035, J037, J038, J042, J044-J059, J061, J063, J064
exp29	Informal	VAT taxed	Food	Gastodiario	
exp30	Informal	VAT taxed	Other non-food goods and services		Clave= C001-C024, D001-D026, E009-E012, E016, E018, E019, E023-E025, E027, E028, E030-E033, F007, F013-F017, G008-G010, G020-G022, H001-H119, I001-I026, K001-K044, L001-L028, M002-M018, N001-N005, N008-N010, T901, T903, T904, T906-T909, T911-T915
exp31	Informal	VAT taxed	Health-related goods and services		Clave= J002, J003, J006, J011, J012, J015, J019, J040, J041, J043, J060, J065-J071
exp32	Informal	VAT taxed	Petrol		Clave= F010,F011,F012
exp33	Informal	VAT taxed	Telecoms		Clave= F001 , F002, F003, F004, F005, F006, F008, F009
exp34	Informal	VAT and duties	under 14° alcohol	Gastodiario	Clave=A228, A231 ,A234, A232, A238
exp35	Informal	VAT and duties	14°-20° alcohol	Gastodiario	Clave= A226, A237
exp36	Informal	VAT and duties	over 20° alcohol	Gastodiario	Clave = A223, A225, A227, A229, A230, A233, A235, A236
exp37	Informal	VAT and duties	beer	Gastodiario	Clave= A224
exp38	Informal	VAT and duties	tobacco	Gastodiario	Clave= A239, A240

Adult file

Variables	Formal/ Informal	Gross	Definition	ENIGH 2008 file	Variable description
folioviv			Residence identifier		
foliohog			Household identifier		
numren			Individual identifier	pobla08	
age			Age	pobla08	
sex			Sex	pobla08	
education			Highest qualification achieved	pobla08	
employment status			Whether in work	pobla08	
inc1	formal	Gross	Wages and salaries	ingresos, pobla08, trabajos	clave=P001, P002,P003, P006, P011, P015
inc2	formal	Gross	Overtime	ingresos, pobla08, trabajos	clave= P004
inc3	formal	Gross	End-of-the-year bonus	ingresos, pobla08, trabajos	clave=P009
inc4	formal	Gross	Incentives, rewards and prizes	ingresos, pobla08, trabajos	clave=P005
inc5	formal	Gross	Holiday bonuses and allowances in cash	ingresos, pobla08, trabajos	clave=P007
inc6	formal	Gross	Profit sharing from secondary subordinated work and end-of-the-year bonus	ingresos, pobla08, trabajos	clave=P008, P019
inc7	formal	Gross	Income from pensions	ingresos, pobla08, trabajos	clave=P032
Inc8	formal	Gross	Other labor income	ingresos, pobla08, trabajos	clave=P013, P017, P018
inc9	formal	Net	Non labor income (unearned)	ingresos, pobla08, trabajos	clave=P012, P016, P023- P031, P033- P080
inc10	informal	Gross	wages and salaries	ingresos, pobla08,	clave=P001, P002,P003,

				trabajos	P006, P011, P015
inc11	informal	Gross	Overtime	ingresos, pobla08, trabajos	clave= P004
inc12	informal	Gross	End-of-the-year bonus	ingresos, pobla08, trabajos	clave=P009
Inc13	informal	Gross	Incentives, rewards and prizes	ingresos, pobla08, trabajos	clave=P005
Inc14	informal	Gross	Holiday bonuses and allowances in cash	ingresos, pobla08, trabajos	clave=P007
Inc15	informal	Gross	Profit sharing from secondary subordinated work and end-of-the-year bonus	ingresos, pobla08, trabajos	clave=P008, P019
Inc16	informal	Gross	Income from pensions	ingresos, pobla08, trabajos	clave=P032
Inc17	informal	Gross	Other labor income	ingresos, pobla08, trabajos	clave=P013, P017, P018
Inc18	informal	Gross	Non labor income (unearned)	ingresos, pobla08, trabajos	clave=P012, P016, P023- P031, P033- P080

A key aspect of this project is the use of alternative assumptions about informality, evasion and missing data to test the sensitivity of results to our baseline assumptions. In order to do this, there will be separate copies of the input data files embodying the different assumptions (e.g. under-reporting income by a constant factor versus assuming it differs across the income distribution).

For example, we will change the definition of formal workers to be those receiving any of the following social benefits: IMSS, ISSSTE, state ISSSTE or PEMEX, not necessarily through work. Given that some individuals are covered by more than one system, testing our results to alternative classification of formal workers will also change their calculated gross income to the extent that the contributions to social security differ across systems. This will be reflected in the adult file. This will change the tax base for the income tax calculation and hence the distributional analysis. We will also test our results to the assumption that all formal workers pay taxes on all its income. To correct in different ways for under-reporting income we will generate different household and adult files containing different amounts for total income and labor and non-labor income. See section 3.5 for a detailed description of how we plan to introduce these different assumptions.

APPENDIX D: DESCRIPTIVE STATISTICS

Table D.1 Correlations between Income Measures and Expenditure Measures

Variables being compared	Correlation Coefficient		
	Whole Sample	Above Median	Below Median
<i>Total Income:</i>			
Per capita and OECD Equivalence Scale	0.9774	0.9779	0.9279
Per capita and 80-55 Equivalence Scale	0.9912	0.9913	0.9738
Total Income and Monetary Income	0.9732	0.9733	0.8601
<i>Total Expenditure:</i>			
Per capita and OECD Equivalence Scale	0.9595	0.9537	0.9482
Per capita and 80-55 Equivalence Scale	0.9872	0.9853	0.9813
Total Expenditure and Monetary Expenditure	0.8035	0.7688	0.8148
<i>Income and Expenditure</i>			
Total Income and Total Expenditure	0.4212	0.3805	0.5820
Monetary Income and Monetary Expenditure	0.3567	0.3245	0.5258

Source: ENIGH (2008) concentrado file, unadjusted.

Table D.2 Informality in the labor market

Informality rate	59.75%
% of annual net labor income accounted for by informal workers	41.99%
% of annual gross labor income accounted for by informal workers	38.83%

Notes: Figures are calculated using sampling weights (variable 'factor' in ENIGH 2008). Workers are those that declared a positive amount of (net) income through labor. Formal workers are defined as those receiving any of the following social benefits through work: IMSS, ISSSTE, state ISSSTE or PEMEX. Gross labor income is calculated using as a base the CIEP simulator (2010). Income is adjusted using the Altimir factor.

Source: ENIGH 2008 and CIEP (2010)

Table D.3 Comparison of prices in the formal and informal sector, for taxed goods.

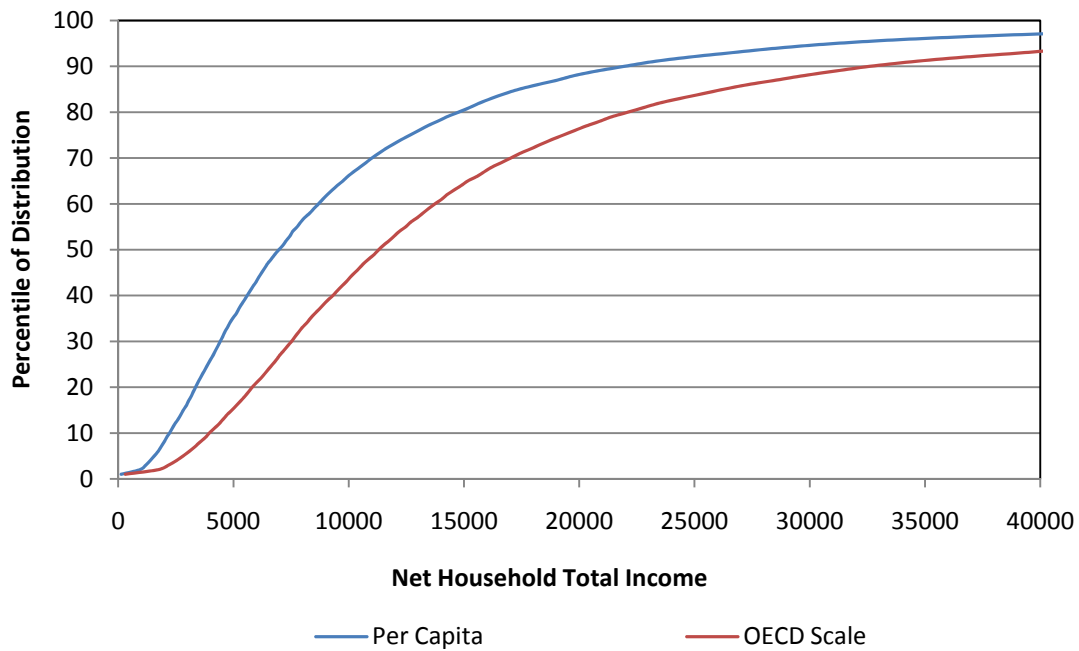
Good	Formal Price (Mean)	Informal Price (Mean)
Cooked Chicken	\$49.58	\$43.44
Other Food	\$43.36	\$39.19
Soft Drinks	\$8.38	\$9.44

Notes: Informal expenditure is defined as those goods and services purchased from informal vendors such as street markets and stalls (ENIGH variable lug_com equal to 1, 2, or 3). Prices are unit values, calculated as the ratio between the variable 'gasto' and 'cantidad' in the ENIGH 2008 file 'gastodiario'. Cooked chicken corresponds to clave equals "A200", Other food corresponds to clave equals "A202", and soft drinks to "A220". The classification of goods and services into the different taxation categories corresponds to the VAT and Duties systems valid in 2008; it is based in CEFP (2009b)..

Monthly expenditure figures are shown.

Source: ENIGH 2008

Figure D.1 Cumulative Frequency of Distribution of Household Net Total Income



Source: ENIGH (2008) concentrado file, unadjusted. Same source for D.1 to D.6.

Figure D.2 Cumulative Frequency of Distribution of Household Net Monetary Income

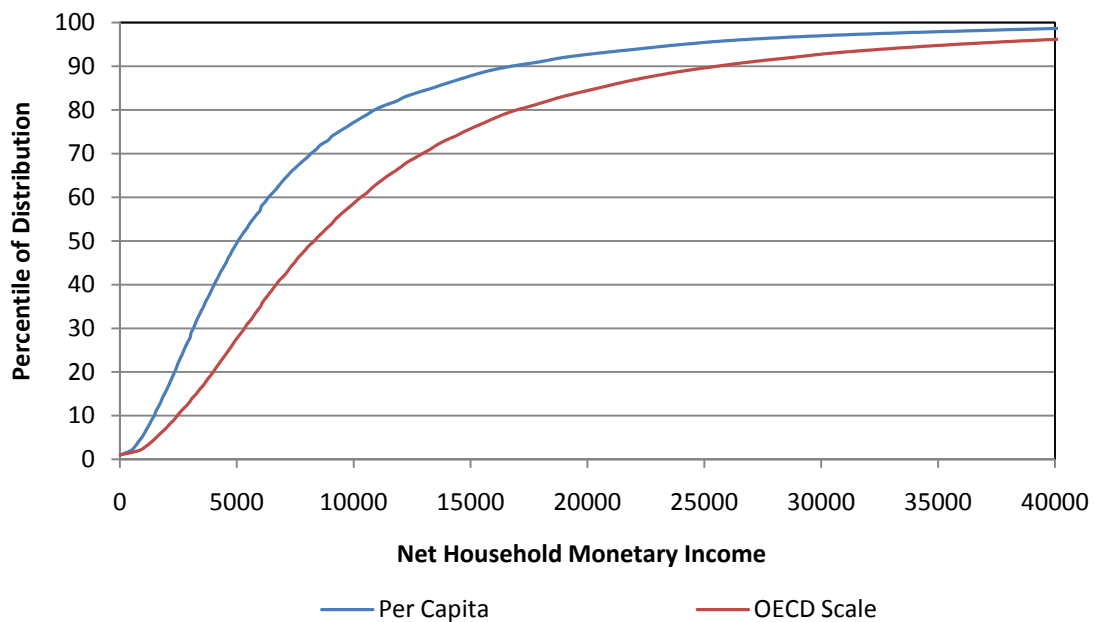


Figure D.3 Cumulative Frequency of Distribution of Total Household Expenditure

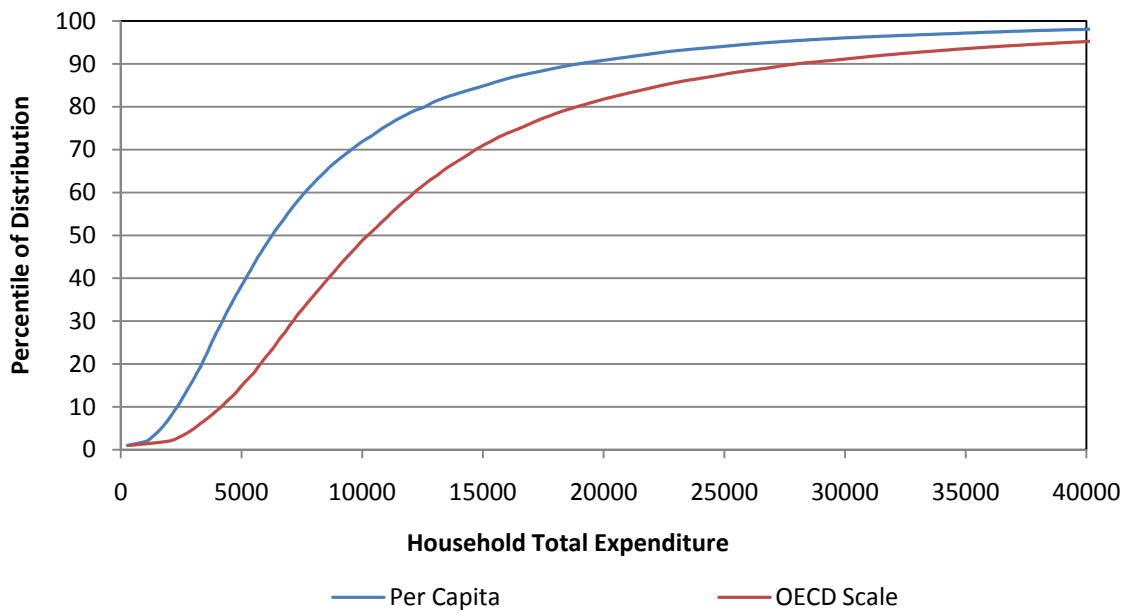


Figure D.4 Cumulative Frequency of Distribution of Monetary Household Expenditure

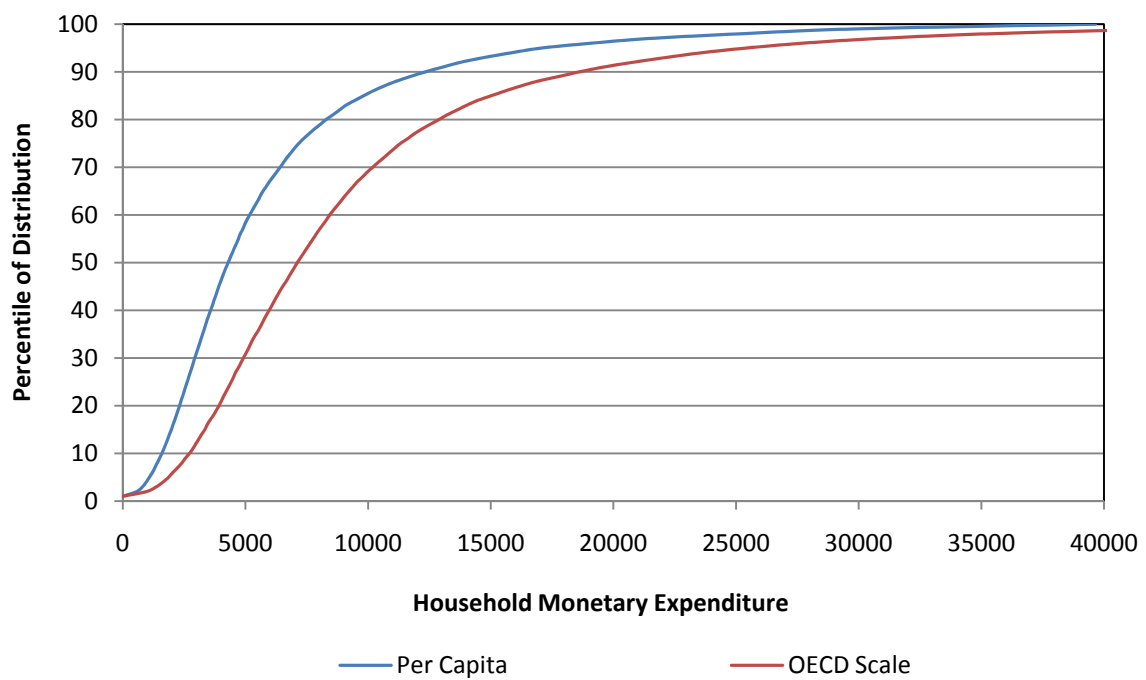


Figure D.5 Distributions of Expenditure and Income

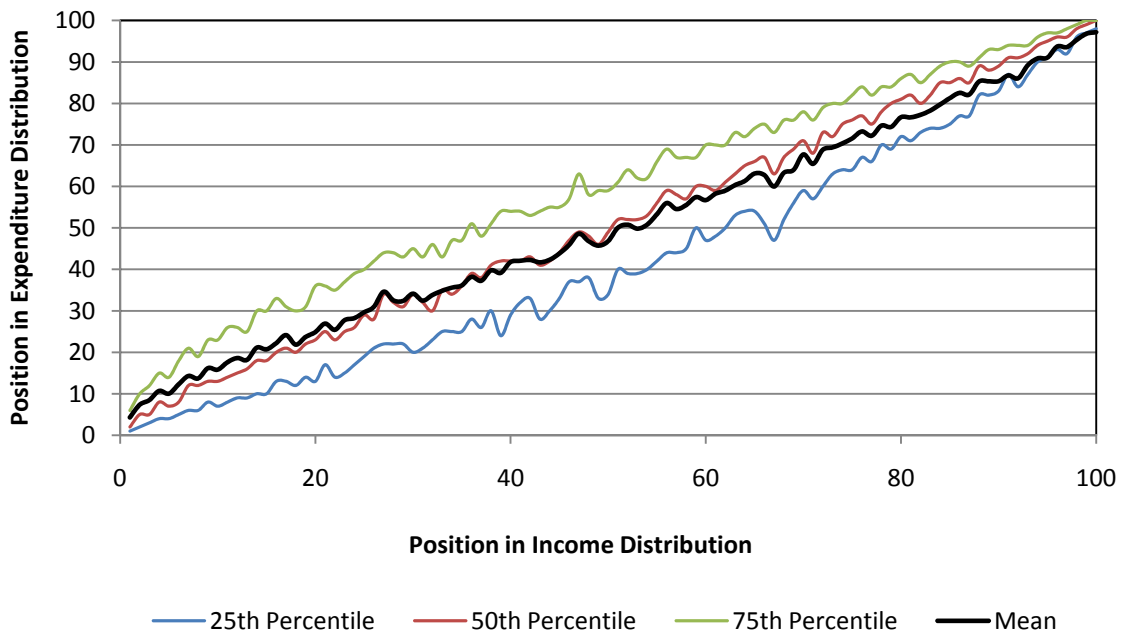


Figure D.6 Distributions of Expenditure and Income

