Is the gasoline tax regressive in the twenty-first century? Taking wealth into account

Jordi Teixidó (EUI)
Stefano Verde (EUI)

Abstract

Poterba (1991a) has much influenced the literature on the distributional effects of carbon pricing. Poterba argues that the incidence of energy/environmental taxes across households is better appreciated if the relative tax burdens are measured against total expenditure, interpreted as a proxy for lifetime income, instead of annual income. This way, however, since the distribution of total expenditure is structurally more uniform, the incidence of energy price increases is always less regressive than when annual income is used. This outcome is often taken to lessen the relevance of equity concerns regarding carbon pricing. Almost twenty-five years after Poterba (1991a), Piketty (2014) revived the idea that wealth is a dimension of economic welfare constituting an increasingly important source of inequality. We show that omitting wealth in measuring ability to pay means underestimating the regressivity of carbon pricing and its inequity towards younger people. Using household-level data and statistical matching, we revisit Poterba's application and compare the distributional incidence of the US gasoline tax for different measures of ability to pay: total expenditure, income and wealth-adjusted income. Regressivity is not a reason to forgo carbon pricing as a cost-effective approach to climate mitigation, but calls for consideration and compensation of the distributional effects.

1. Introduction

The idea of taxing fossil fuels in proportion to their carbon content goes back as far as the 1970s, when the threat of anthropogenic climate change started to be recognized[1]. In 1990, Finland was the first country to introduce a carbon tax, followed shortly after by the Netherlands, Sweden and Norway. Today carbon pricing, whether in the form of carbon taxes or cap-and-trade systems, is in force in several countries, but overall is far from being sufficiently diffuse or deep to significantly improve the prospects of climate change. Global greenhouse gas (GHG) emissions have been rising steadily since the industrial revolution and will continue to do so unless counteracting policies are ramped up. In this respect, a change of gear seems finally in sight. An intensification of mitigation policies around the world should materialize under the framework set out by the Paris Agreement[2]. Accordingly, in the next few years carbon pricing is expected to become more widespread and deeper than it is currently.

Most economists favour carbon pricing in that it is a cost-effective approach to reducing GHG emissions (Baumol and Oates, 1971). Nevertheless, carbon pricing in the real world is not popular or easy to implement. For carbon pricing to be politically sustainable, its side effects need to be effectively managed[3]. By raising the cost of energy, unilateral carbon pricing can be detrimental to the international competitiveness of domestic energy-intensive firms. At the same time, carbon pricing tends to affect the poor more than the wealthy in relative terms. That is, it tends to be regressive, at least in developed economies[4]. The revenues generated by carbon pricing, be they the yield of a carbon tax or of the auctions of emission allowances under a cap-and-trade system, could be used to at least partially offset these undesirable effects. Though this is easier said than done[5], the deeper the level of carbon pricing, the more critical it is that both the competitiveness and distributional issues are properly addressed.

This paper offers a new perspective and new empirical evidence on the distributional incidence of gasoline taxes and, by extension, of carbon pricing across households. Specifically, it fills a gap in the literature by considering wealth (net worth) as a dimension of economic welfare additional to income. This innovation provides us with a more accurate representation of reality, in which the wealth owned by a person, or a household, contributes to her ability to pay (taxes). In this sense, ignoring wealth is an omission that alters the portrait of distributional effects, because wealth is both

---

[1] See, for example, the early contributions of Nordhaus (1977a, 1977b), among the first proponents of carbon taxation.
[2] The Paris Agreement is the international agreement, under the United Nations Framework Convention on Climate Change, dealing with climate mitigation, adaptation and finance, starting in the year 2020.
[3] Moreover, a growing literature deals with the public's cognitive difficulties and worldviews that hinder its adoption. Drews and van den Bergh (2015) provide a comprehensive literature review on the determinants of public support for climate policies.
[4] In developed economies, the income elasticity of energy demand is typically smaller than 1. The same is not necessarily true for developing economies, given the different structure of household demand.
[5] Earmarking is somewhat infrequent and unpopular among economists, as it generally means foregoing alternative more efficient uses of the revenues.

more concentrated than income and also imperfectly correlated with it. This issue appears to be increasingly relevant in light of Thomas Piketty's warning, in his *Capital in the twenty-first century* (Piketty, 2014), that wealth concentrations have been rising and may well continue to rise unless corrective policies are undertaken.

While taking wealth into account is generally desirable for the completeness of any equity assessment, it is particularly opportune in relation to carbon pricing. This is the case for different reasons. First, carbon pricing without a redistributive mechanism linked to it effectively amounts to financing a public good, namely climate stability, through regressive taxation. Not surprisingly, it often encounters strong resistance motivated by equity concerns. Second, the need to reduce GHG emissions and the related commitment of the Paris Agreement, suggest that carbon pricing will become deeper in the near future. Third, following James Poterba's work in this field (1989, 1991a, 1991b), a significant proportion of the literature plays down the relevance of the distributional effects of carbon pricing. This outcome stems from specific methodological choices, notably that of considering (expected) lifetime ability to pay instead of (observable) current ability to pay.

Using household-level data from the 2012 round of the US Consumer Expenditure Survey (CE) and from the 2013 Survey of Consumer Finances (SCF), we revisit Poterba's 1991 seminal paper *Is the gasoline tax regressive?* (Poterba, 1991a). Poterba's analysis is extended, empirically, by imputing observed wealth in the SCF to households in the CE and, theoretically, by considering wealth as one dimension of economic welfare and, hence, as a complementary measure of ability to pay. Based on annual gasoline expenditure, we estimate the economic burden of the federal gasoline tax ($0.184/gallon) relative to three alternative measures of ability to pay: *a)* annual total expenditure, as a proxy for lifetime income (Poterba's approach), *b)* annual income and *c)* annual wealth-adjusted income, which is annual income augmented with a wealth annuity and imputed rental income (for home owners). The analysis of the results consists in the comparison of the three measurements of the relative tax burdens, first, across the respective distributions of ability to pay measures and, then, across the distribution of the head of household's age. The positive correlation between wealth and age, due to the first accumulating over time, indeed implies that the distributional incidence of carbon pricing across age groups changes depending on whether wealth is considered. Considerations about intergenerational equity are generally relevant to climate policy given the difference between the young and the elderly in both the responsibilities for causing climate change and the related costs faced in prospect.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature. Section 3 explains why wealth should be considered in this context. Section 4 derives and compares the distributional incidence of the US federal gasoline tax according to alternative measures of ability to pay. Section 5 concludes.

2. Literature review

The connections between gasoline taxation and carbon pricing are such that our analysis while dealing with the former can be relevant also to the latter. Focusing on gasoline taxes simplifies the analysis in terms of data availability, methodology and assumptions, while remaining sufficient to highlight the role of wealth in the equity assessment of any policies affecting energy prices.

Apart from the substitution between motor fuels with different carbon content (principally gasoline and auto diesel), studying the economic effects of gasoline taxes is effectively equivalent to studying the effects of carbon pricing in the road transportation sector. A second connection between gasoline taxes and carbon pricing concerns the relative degree of regressivity. Price increases in motor fuels are typically less regressive than price increases in home fuels (principally electricity and natural gas), as the demand for the first is more income elastic than that for the second (e.g., Barker and Köhler, 1998; Tiezzi, 2005; Callan *et al.*, 2009; Ekins *et al.*, 2011; Hassett *et al.*, 2012; Kosonen, 2014; Flues and Thomas, 2015; Verde and Pazienza, 2016). As a result, gasoline taxes are usually less regressive than carbon pricing when this is operating in sectors of the economy other than transportation, notably electricity generation and the residential sector.

The following literature review focuses on the methodological aspects most relevant to our analysis. It first covers the empirical studies on gasoline taxes and, subsequently, those on carbon pricing.

2.1 The distributional incidence of gasoline taxes

The empirical literature on the distributional incidence of gasoline taxes largely uses household survey data to estimate tax burdens, usually quantified by tax payments (or welfare changes, when price changes are considered within demand systems), across income levels or socio-demographic characteristics. The frameworks used are either static or allow for demand response to price changes, sometimes within demand system models estimated under separable utility assumptions. In the applications to developed economies, gasoline taxes are found to be regressive to varying degrees or approximately proportional, in this case often with middle-income households bearing the heaviest burdens[6]. Importantly, however, the results are not independent from some methodological choices. As noted by Sterner (2012a), at least two types of choice can affect the distributional outcome significantly. One concerns the inclusion or exclusion of the households that do not own any vehicles. Since most of these households are at the bottom of the income distribution, their inclusion (exclusion) in the calculations results in a less (more) regressive outcome. The second choice concerns the variable measuring the ability to pay or, rather, the time horizon over which the ability to pay is valued. This is typically the present or, in an ex-ante perspective, a person's lifetime. The longer the time horizon, the less variable is the distribution of economic welfare, due to both earnings

[6] In developing economies, gasoline taxes are generally progressive (Sterner, 2012a).

patterns over time and income mobility, so gasoline taxes are less regressive over a lifetime. For a number of countries, Sterner (2012b) contrasts the different distributional incidence of the same gasoline taxes obtained using the current ability to pay approach and the lifetime approach.

The present paper deals with the implications of the second choice above. In addressing this question, we are not the first to take a critical stance: Chernick and Reschovsky (1992, 1997, and 2000) were the first, but also the last as far as we are aware. They brought arguments and evidence that fundamentally question James Poterba's lifetime approach to estimating the distributional incidence of gasoline taxes (Poterba, 1991a) and carbon taxes (Poterba, 1991b). Poterba's approach, which leads to the conclusion that these taxes are not regressive over a lifetime, consists in the use of current total expenditure as a proxy for lifetime income and, therefore, as a measure of lifetime ability to pay. Chernick and Reschovsky point out that this approach, which emanates from Milton Friedman's permanent income theory of consumption (Friedman, 1957) and the companion life-cycle model of saving (Ando and Modigliani, 1963), rests on a set of very strong assumptions, namely: *a)* income mobility is very high; *b)* gasoline consumption decisions are made on the basis of lifetime income; and *c)* total consumption is a constant fraction of lifetime income. Using longitudinal data, they cross-check Poterba's results by deriving the distributional incidence of the US gasoline tax over an 11-year period, finding that, with the exception of the bottom 11-year average income decile, the incidence is in fact only slightly less regressive than when annual income is used. The authors emphasize that the main reason for the similarity between annual and intermediate-run tax burdens is low income mobility. Thus, the volatility of annual income as an argument in favour of the lifetime approach appears to be justified only for the lowest levels of annual income.

In spite of Chernick and Reschovsky's analysis and findings, many subsequent studies assess the distributional incidence of gasoline taxes and carbon taxes using Poterba's lifetime approach. Only few adopt the lifetime perspective while applying more sophisticated approaches than Poterba's, including notably Fullerton and Rogers (1993) and Bull *et al.* (1994). Both the frequent lack of income data (or of sufficiently good quality income data) in household surveys and its computational simplicity, may at least partly explain the fortune of Poterba's approach as reflected in the number of its followers.

2.2 The distributional incidence of carbon pricing

The literature on the distributional effects of carbon pricing is methodologically more diverse than that on gasoline taxes. This is the case because carbon pricing can cover an area of the economy that is broader than the transportation sector. Accordingly, economy-wide models are often used: usually either computable general equilibrium (CGE) models or macroeconometric models (occasionally combined with microsimulation models). The advantage of using these models for distributional analysis is that secondary and general equilibrium effects are taken into account. The CGE literature, in particular, stresses the capability of these models to capture the distributional effects of carbon pricing occurring through both the uses side of income, i.e. consumption and savings, as well as the sources side of income, i.e. the returns to labour and capital. The distributional incidence of carbon pricing is thus given by the sum of the effects unfolding over the two sides of income, which in turn depend on the use of the revenues generated by carbon pricing ("revenue recycling") and the (related) impact on the economy. Over time, progressive sources-of-income effects may partially or even entirely offset the regressive uses-of-income effects typically captured by partial equilibrium models. Rausch *et al.* (2011) and Dissou and Siddiqui (2015) illustrate this type of CGE result.

The CGE literature also stresses the efficiency-equity trade-off between alternative uses of the revenues generated by carbon pricing. Namely, the redistributive options can tackle the regressive distributional effects, but not the efficiency loss of the economy due to carbon pricing. Vice versa, through the reduction of distortionary taxes, typically on labour or capital, the efficiency revenue recycling options tackle the economy's efficiency loss, but not the regressive distributional effects. This trade-off relates to our analysis in that underestimating regressivity makes efficiency-enhancing tax cuts unduly more attractive than the redistributive alternatives.

3. Why considering wealth

In official statistics, annual income is the standard measure of ability to pay used to determine the degree of tax progressivity or regressivity. However, we have seen that alternative measures of ability to pay are considered in the empirical literature. Notably, estimated lifetime income measuring lifetime ability to pay is often used to determine the lifetime distributional incidence of gasoline taxes or carbon pricing. Fullerton and Rogers (1991), who are among the most prominent authors considering the lifetime perspective, argue that policymakers should be concerned with "short run equity" as well as "long run equity". In their words, "the fairness of a tax should be evaluated both on how current taxes reflect current ability to pay and on how lifetime taxes reflect lifetime ability to pay".

The central argument of the present paper is that considering wealth as a complementary measure of ability to pay constitutes an improvement on the use of sole annual income and, all the more so, of lifetime income. In our view, which seems to contrast with Fullerton and Rogers' above statement, short run equity and long run equity are not equivalent or equally relevant. The two concepts fundamentally differ in that the first is observable while the second can only be predicted. As far as equity judgments are concerned, realised outcomes matter, while predicted outcomes do not matter as much. Secondly, lifetime approaches necessarily rest on sets of assumptions which affect the reliability of the results. Nonetheless, the same lifetime results are often presented as lessening the relevance of equity concerns regarding carbon pricing (e.g., Hassett *et al*., 2009; Sterner, 2012a, 2012b; Kosonen, 2014; Mathur and Morris, 2014; Parry, 2015; Williams, 2016). Crucially, this takes us a step away from the reality of the equity problem. There is in fact greater urgency for the measurement of ability to pay to be extended "in perimeter", by considering wealth, rather than in time, as with the lifetime perspective. The present section elaborates on these points.

3.1 Ex post equity versus ex ante equity (or short run equity versus long run equity)

The lifetime perspective in evaluating the distributional incidence of gasoline taxes and carbon pricing, and of all taxes in general, implies that interpersonal comparisons are based on expected lifetime ability to pay as opposed to observed current ability to pay. Yet, because expected outcomes may obviously not coincide with subsequently realised outcomes, people normally make equity judgments based on observed, realised welfare differentials. For the same reason, welfare programs are calibrated based on observed welfare differentials, not expected ones. As Warren (1980) points out, expectations are central to the economic theory concerned with the making of rational choices ex ante; but fairness in taxation should depend – and indeed does depend, in the real world – on realised outcomes, not expectations.

We thus find the lifetime perspective of interest for the analytical insights that it offers, but not as much for the utility of the related policy implications, notably for informing a measure of ability to pay[7]. By contrast, considering wealth in measuring current ability to pay is an innovation that provides us with a better representation of reality. This is because wealth is a dimension of economic welfare (see below) and people – we can safely assume – internalize observed wealth differentials (just like income or consumption differentials) in making equity judgments. Nevertheless, as it stands, the literature on the distributional incidence of gasoline taxes and carbon pricing ignores wealth altogether.

3.2 Wealth as a dimension of economic welfare

Which, among income, consumption and wealth, should be targeted by direct taxation is a question long debated by economists. The matter is complex because it relates to philosophical views as well as both economic and practical considerations. Related to this question, there now seems to be general agreement that income, consumption and wealth capture different dimensions of a person's economic welfare. The Commission on the Measurement of Economic Performance and Social Progress, a.k.a. the Stiglitz-Sen-Fitoussi (SSF) Commission, recommended in its final report that income, consumption and wealth be considered together to measure economic welfare and, therefore, to measure ability to pay (Stiglitz *et al.*, 2009). In the same report, the rationale for the use of the three indicators is explained as follows:

---

[7] Though more for practical than for philosophical reasons, Fullerton and Rogers (1991) concede that "the lack of savings data and the complexity involved in simulating such data may make lifetime incidence more of an academic exercise than an operational policy tool."

"Income flows are an important gauge for the standard of living, but in the end it is consumption and consumption possibilities over time that matter. The time dimension brings in wealth. A low-income household with above-average wealth is better off than a low income household without wealth. The existence of wealth is also one reason why income and consumption are not necessarily equal: for a given income, consumption can be raised by running down assets or by increasing debt, and consumption can be reduced by saving and adding to assets. For this reason, wealth is an important indicator of the sustainability of actual consumption."

About forty-years before the SFF Commission, Weisbrod and Hansen (1968) were the first to study the implications of considering wealth (net worth) as a store of potential consumption and, therefore, of economic welfare. The authors devised a method whereby income and wealth are combined into a single indicator of economic welfare. They then explored the implications of using the income-wealth indicator ("wealth-adjusted income", as we call it below) for the assessment of economic inequality, including tax progressivity and regressivity, and for the prediction of consumption behaviour. The key element of the authors' analysis is the imperfect correlation between income and wealth, which means that households' welfare ranking is different depending on whether income or the income-wealth indicator is used. Other studies have subsequently dealt with the same idea, including Taussig (1973), Wolfson (1979) and Radner and Vaughan (1987). More recently, Weisbrod and Hansen's income-wealth indicator was refined and integrated in the Levy Institute Measure of Economic Well-being (LIMEW)[8], from which our analysis below borrows several methodological aspects. Applications of the LIMEW indicator include, among others, Wolff *et al.* (2005) and Wolff and Zacharias (2007, 2009).

3.3 Wealth inequality and carbon pricing: "the elephant in the room"

It is a well-known fact that the distribution of income and the distribution of wealth significantly differ one from the other, the latter being more concentrated than the former. In *Capital in the twenty-first century*, Piketty (2014) examines the evolution of the two distributions, primarily in Europe and in the US, since before the nineteenth century. While wealth concentrations are much lower today compared to the peak in the years preceding World War I, one of Piketty's main conclusions is that very high wealth concentrations may characterise the economy of the twenty-first century. The last four decades have indeed seen a positive trend in wealth concentrations over time, especially in the US, which may well continue if certain conditions persist. In this context, taking wealth into account, for evaluating people's economic welfare and addressing distributional issues, is all the more desirable. Yet, to date annual income remains the only measure used for such purposes.

[8] http://www.levyinstitute.org/research/the-levy-institute-measure-of-economic-well-being

The utility of considering wealth is general, in the sense that it would benefit any type of assessment concerning economic equity. However, we deem it to be particularly relevant for appreciating and, thus, for dealing with the opposition to carbon pricing motivated by equity concerns. First, one needs to recall that the ultimate purpose of carbon pricing is to maintain a stable climate, which is a (global) public good. Second, presumably everyone supports public goods as long as their cost is shared in a way that is perceived as fair. Indeed, it is difficult to imagine that no opposition would arise to the financing of a public good through regressive taxation. Though never clearly acknowledged, as far as we are aware, carbon pricing without a redistributive mechanism linked to it effectively corresponds to this type of setting [9]. Considering wealth in evaluating economic welfare is therefore all the more desirable in relation to carbon pricing. Adapting a well-known figurative expression, if wealth is the elephant in the room that somehow goes unnoticed, climate change as a public good problem makes the room smaller: so the elephant is even bigger in relative terms.

4. The distributional incidence of the US gasoline tax

In the US, three tax layers apply to the consumption of gasoline and auto diesel, namely, federal taxes, State taxes and local taxes. The federal tax rate on gasoline is currently 0.184 $/gallon and has not changed since 2006. The federal tax rate on auto diesel is 0.244 $/gallon. State and local taxes can differ significantly across the country and, as compared with the federal taxes, are more frequently subject to revisions. In recent years, growing concerns related to declining fiscal revenues and high CO_2 emissions meant that the option of raising gasoline taxes, which are very low compared to those in most developed economies, has received increasing consideration in the American public policy debate. Still, raising gasoline taxes is anything but a popular measure.

Using data referring to the year 2012, we analyse the distributional incidence of the US federal gasoline tax across households. The purpose of the study is to show the implications of using different ability to pay measures for the resulting distributional incidence. Our contribution is bringing wealth (net worth) into the equation as a dimension of economic welfare complementary to income. This, we argue, offers a more accurate rendering of the distributional incidence. Effectively, we revisit and extend Poterba's seminal paper *Is the gasoline tax regressive?* (Poterba, 1991a) by introducing a third measure of ability to pay, namely wealth-adjusted income, alternative to both income and total expenditure. Wealth-adjusted income only differs from annual income, which is the standard measure of current ability to pay, in that it also includes the value of potential consumption stored in currently

[9] The provision of public goods is usually financed through the general tax system, which in the modern fiscal state is not regressive.

owned wealth (see Section 4.1.2 below). It is thus a more comprehensive measure of current ability to pay[10].

The first part of this section is devoted to *a)* the data and the data work for imputing wealth to the households in our sample, and *b)* the definitions and the assumptions made for determining wealth-adjusted income. The second part deals with the differences in distribution between the alternative ability to pay measures. The third part examines the respective differences in distributional incidence both across welfare levels and the head of household's age.

4.1 Data

Our analysis is based on household-level data from the 2012 round of the US Consumption Expenditure Survey (CE). Our sample consists of 2,179 households, those for whom annual expenditure could be derived[11]. Developed by the Bureau of Labor Statistics (BLS), the CE is the most comprehensive data source on US households' consumption choices, including information on expenditure, income and socio-demographics[12]. The CE serves well the purpose of our study, just as for most of the closely related US literature.

Crucially, however, the CE does not contain information (or, rather, not sufficiently accurate information) on households' wealth. To overcome this limitation, we use statistical matching (a.k.a. data fusion) whereby household-level information on wealth is imported from the US Survey of Consumer Finances (SCF) into our CE sample. After performing the statistical matching, we follow Wolff *et al.* (2005) and Wolff and Zacharias (2007, 2009) in developing an indicator of economic well-being which aptly combines households' income and wealth. This indicator, measuring what is referred to as wealth-adjusted income, allows us to assess the distributional incidence of the gasoline tax while taking the wealth dimension of a household's ability to pay into account. The three measures of ability to pay, namely income, total expenditure and wealth-adjusted income, are flow variables directly comparable to one another.

4.1.1 Statistical matching

The purpose of statistical matching is to obtain joint information on the not jointly observed variables (D'Orazio *et al*., 2006). The most common setting is that of two surveys drawn from the same population and sharing a set of common variables, **X**, typically socio-demographic variables, but not other variables, Y and Z, whose relationship is of interest. In practice, we use statistical matching to assign specific households' observed wealth in the SCF sample to the households in our CE sample.

[10] Using wealth-adjusted income as a measure of ability to pay should partly alleviate the issue of temporarily very low annual incomes that make the gasoline tax, or carbon pricing for that matter, look unduly more regressive (Chernick and Reschovsky, 1992, 1997, and 2000). This is because, for example, pensioners who may report very low income may have accumulated some wealth.

[11] In the 2012 round of the CE, this is the number of households with four quarterly interviews.

[12] The BLS uses the CE to periodically revise the expenditure weights of the Consumer Price Index.

The imputation of wealth is based on the relationships between the common variables (**X**) and both wealth (Y) and gasoline expenditure (Z). The resulting fused dataset is the initial CE sample (which includes information on gasoline expenditure) augmented with imputed wealth.

After harmonizing the variables shared by the two surveys, only those with similar empirical distributions in the two datasets and, also, statistically associated with both wealth and gasoline expenditure in the donor and recipient datasets, respectively, are selected as matching variables (D'Orazio *et al*., 2006). These turn out to be the following: household income, housing tenure, age of the reference person, her education level, her marital status and her employer type. Propensity score matching (Rässler, 2002) is then used to assign observed wealth in the SCF to each household in the CE dataset. Propensity scores are derived based on the matching variables above and the Mahalonobis distance function is applied to pair households across the two datasets. The resulting fused dataset satisfactorily meets the standard validity requirements of statistical matching (Rässler, 2002, 2004). The details on the matching procedure are provided in Section A of the Appendix.

4.1.2 Measuring wealth-adjusted income

Following Wolff *et al.* (2005) and Wolff and Zacharias (2007, 2009), we derive a measure of annual wealth-adjusted income that uses annual money income (earnings, property income, and other money income) as its basis, subtracts property income to avoid double counting (as explained below), and then adds a constant wealth annuity as well as imputed rental income for home owners. Wealth is net worth, namely the current value of all marketable or fungible assets less the current value of all debts. Only assets that can be readily converted to cash and so into potential consumption, without compromising current consumption, are considered. Accordingly, consumer durable goods, future social security benefits and future retirement benefits from defined-benefit private pensions are not included. Table 1 shows summary statistics of households' (imputed) assets and liabilities (per adult equivalent[13]) in the fused dataset[14].

[TABLE 1]

To combine wealth and income into a single ability to pay measure, wealth as just defined needs to be converted into a flow variable. It is here converted into a stream of constant annual payments (annuities) covering the expected remaining life of the head of household or of the younger spouse if there is one[15]. Basically, the sum of current net worth and the relative expected future incomes is spread evenly over time and exhausted at the end of the expected lifetime. The wealth

[13] The new OECD equivalence scale is used, in which the head of household weighs 1, all other household members aged over 13 weigh 0.5 each, and those under 14 weigh 0.3 each.

[14] Sampling weights are applied in all the calculations presented in this paper (including summary statistics).

[15] Together with age, we take into account gender differences in life expectancy. Life expectancy estimates are taken from Arias (2015).

annuity is derived based on a weighted average of historical rates of return on different types of assets. Specifically, we use the average return rates indicated in Wolff and Zacharias (2009), updated to the period 1972-2012[16] (see Table B1, in the Appendix). As these rates already include both capital gains (realized and unrealized) and any income the assets may generate, reported property income (interests, dividends and rents) is subtracted from household money income to avoid double counting[17].

Furthermore, as housing is a universal need, home ownership frees the owner from the obligation of paying a rent, leaving an equivalent amount of financial resources for other uses. Again, following the LIMEW approach, our measure of wealth-adjusted income takes this factor into account. The rental income imputed to the households owning their principal residence is calculated by multiplying the value of the dwelling by the US average ratio of (imputed) rent-to-home value for owner-occupied homes, which was 5.7% in 2012[18]. The cost of owning the dwelling is included in net worth, which accounts for the outstanding mortgage debt[19].

Formally, wealth-adjusted income (WI) of household *h* is calculated as follows:

$$WI_h = MI_h - PI_h + WA_h + IRI_h \tag{1}$$

where MI is money income, PI is property income, WA is the wealth annuity, and IRI is imputed rental income (different from zero only for home owners).

$$WA_h = \frac{\sum_{k=1}^{K=5} A_{k,h}(1+r_k)^{LifeExp_h} + \sum_{j=1}^{J=2} D_{j,h}\left(1+g_j\right)^{LifeExp_h}}{LifeExp_h} \tag{2}$$

where A_k and D_j are the asset and debt aggregates in Table 1, r_k and g_j are the respective return rates (see Table B1, in the Appendix), and LifeExp is expected remaining lifetime (in years).

Table 2 shows the average composition of wealth-adjusted income in the sample used for our analysis.

[TABLE 2]

[16] The data on return rates are from the Federal Reserve's (2012) Flow of Funds Accounts (see Appendix B).

[17] Underreporting of capital income is a well-known issue in household surveys. Capital income is the most underreported type of income in household surveys, with underreporting estimated at up to 50 percent in some OECD countries (Milanovic, 2002).

[18] Information on the sum of imputed rents for US home owners in 2012 is taken from the National Income and Product Accounts, Table 7.12, Line 154 (Bureau of Economic Analysis, 2015). Information on the total value of home owners' principal dwellings in 2012 is taken from the Federal Reserve's (2012) Flow of Funds Accounts.

[19] Table B2, in the Appendix, compares the marginal distribution of our imputed rental income to those of home owners' expected rent if they were to rent out their dwelling and of home renters' paid rent (the CE has information on both). This cross-check indicates that the values of imputed rental income are plausible.

4.2 Differences between the distributions of alternative ability to pay measures

For a household, or an individual, the relative burden of the gasoline tax is here given by the tax payment embedded in her gasoline expenditure relative to her ability to pay. In the literature, the denominator of this ratio is either annual income or annual total expenditure as a proxy for lifetime income. Our contribution is to consider wealth as an additional dimension of economic welfare. However, since wealth is a stock variable, while both income and total expenditure are flows, wealth-adjusted income (rather than wealth) calculated as per above is the third measure of ability to pay directly comparable to the others. Clearly, the distribution of the relative tax burden across households is also dependent on that of the variable at the denominator: the more uneven (dispersed) the distribution of the ability to pay measure, the more uneven the distribution of the relative tax burden too. Moreover, inasmuch as the different measures of ability to pay are imperfectly correlated with one another, the ranking of households by tax burden also depends on which measure of ability to pay is used.

4.2.1 Differences in distribution and households' ranking

With reference to the fused dataset, Table 3 reports descriptive statistics of the sample distributions of *a)* annual income, *b)* annual total expenditure, *c)* wealth (net worth) and *d)* wealth-adjusted income, all of which are expressed in per-adult-equivalent terms. While the median of total expenditure is not much smaller than that of income, the distance widens for the upper parts of the two distributions, as one would expect. The distribution of wealth exhibits negative values up to around the 10^{th} percentile, it becomes positive and rapidly increases thereafter.

[TABLE 3]

The skewness statistics indicate that the distribution of wealth is the most asymmetric, as being more positively skewed than those of income and especially of total expenditure. The kurtosis statistics tell us that the distribution of wealth is also the one with the heaviest tails (relative to the rest of the distribution). We can then deduce that the wealth distribution has a longer right tale. This is reflected in the higher concentration of wealth as measured by the Gini coefficient and pictured in Figure 1.

[FIGURE 1]

As the Gini indices show, despite wealth being highly concentrated, wealth-adjusted income is not much more concentrated than income. As Wolff *et al.* (2009) explain, there are two reasons for the limited difference in terms of concentration between the two variables. First, household income and wealth are not perfectly correlated, so that there are households with low income but high wealth, and households with high income but low wealth. Second, usually, the wealth annuities and imputed rental income together are significantly smaller than annual income (see Table 2). As a result, the inclusion of the wealth annuities in augmented income does not alter the overall distribution of income very much.

Moreover, in principle, alternative ability to pay measures may have equally shaped distributions but entirely different ranking of the statistical units, which are households in our case. In general, the weaker is the correlation between the two variables, the greater is this type of mismatch. Table 4 illustrates the frequency of these changes in households' ranking when switching from one measure of ability to pay to another.

[TABLE 4]

For each of the five pairs of distributions, the rows indicate the shares of total households falling in the same quintiles of the two distributions (in which case the quintile change is equal to 0) or in quintiles that are one to four quintiles apart (in which case the quintile change ranges between -1 and -4 and between +1 and +4). In the first column, the comparison of income vs total expenditure shows that only 46% of all households are equally positioned in the two distributions. The negative values (-1 to -4) correspond to the households that are relatively richer in total expenditure than in income: they add up to 25% of all households; and vice versa for the positive values (1 to 4). The mismatch is slightly more frequent in the comparison of total expenditure vs wealth-adjusted income (second column), while it is clearly less frequent in that of income vs wealth-adjusted income (third column)[20].

4.2.2 Differences across age groups

For the simple reason that people accumulate wealth over time, taking wealth into account in determining ability to pay has a disequalizing effect over households' age dimension. To examine this aspect, we have partitioned our sample into seven groups according to the age of the head of household (Table 5).

[TABLE 5]

The top graph in Figure 2 shows median wealth per adult equivalent by the head of household's age group. The pattern of median wealth across age groups is very clear. The wealth owned by the median household in the top age group, 75-89 years old, is about ten times that of the median household in the 35-44 year-old group. The difference is even more striking if the comparison is made

[20] The mismatch is much more pronounced in the comparisons of both total expenditure and income vs wealth (fourth and fifth column, respectively).

with the two youngest groups; or if wealth cumulated over the three oldest groups is compared with that of the three youngest.

[FIGURE 2]

The bottom graph in the same Figure shows the median values of the different ability to pay measures – income, total expenditure and wealth-adjusted income – by age group. When contrasting income and total expenditure, the most significant differences between the two are observed for the 45-54 and the 55-64 mid groups. The pattern of wealth-adjusted income is such that the distance from income or total expenditure tends to widen with the head of household's age. Thus, while for the youngest households, whether income, total expenditure or wealth-adjusted income is considered does not make much of a difference in absolute terms, it does make a difference for the more mature households. For these households, substantially higher levels of wealth-adjusted income relative to income or total expenditure mean that their ability to pay is significantly underestimated when using one of the two latter measures. Moreover, due again to the highly uneven distribution of wealth across age, ranking effects correlated with age are determined by the inclusion of wealth in the measurement of ability to pay (see Figure C1, in the Appendix).

4.3 The distributional incidence of the gasoline tax by ability to pay measure

We now turn to examining the distributional incidence of the gasoline tax according to the measuring of ability to pay. We first focus on tax regressivity, which concerns the distributional incidence of the tax across levels of ability to pay. We then consider the distributional incidence of the gasoline tax across age groups.

4.3.1 The degree of tax regressivity

For a person or a household, the relative burden of a commodity tax is given by the ratio between the tax payment implicit in her consumption of the good and her ability to pay. The distribution of these burdens across levels of ability to pay determines the degree of regressivity, or progressivity, of the tax. If the tax under study is one already in force, as opposed to a hypothetical new tax or a tax increase (in which cases allowing for demand response is relevant), its degree of regressivity is very well proxied by the distribution of the ratio between expenditure on the given good and ability to pay. The graphs in Figure 3 show median gasoline expenditure as a proportion of the different ability to pay measures, by decile of the same ability to pay variable.

[FIGURE 3]

The median burden of the first decile is clearly highest (11.3%) when ability to pay is measured by income (A graph). The steep decline of the median burden across the income deciles suggests that the gasoline tax is highly regressive. The same conclusion applies when ability to pay is measured by wealth-adjusted income (C graph), but a more rigorous assessment will allow us to ascertain which of the two measures results in a more regressive outcome (see below). By contrast, the distributional incidence of the gasoline tax appears to be significantly less regressive when ability to pay is measured by total expenditure (B graph). As in most of the studies that use total expenditure as a proxy for lifetime income, the gasoline tax is found to be progressive over the lower part of the total expenditure distribution and then to turn to regressive over the better-off deciles.

To quantify the degree of tax regressivity for the three alternative measures of ability to pay, we calculate the Suits index (Suits, 1977). To do this, we first derive each household's tax payment by dividing gasoline expenditure by the relevant average gasoline price[21, 22]. Analogous to the Gini index for its geometrical derivation, the Suits index, S, can take any value between +1 and -1, which correspond to the limiting cases of progressivity (the wealthiest bear the entire tax burden) and regressivity (the poorest bear the entire tax burden), respectively, and is equal to 0 in the case of perfect proportionality. Let y be the cumulative share of overall income, or total expenditure or wealth-adjusted income, and T the cumulative share of overall tax payments,

$$S = 1 - \frac{L}{K} = 1 - \frac{\int_0^{100} T(y)dy}{5000} \tag{3}$$

where L is the area under the Lorenz curve and K the area under the 45-degree line of proportionality ($100 \times 100/2 = 5000$).

We find: $S_I = -0.29$, $S_C = -0.15$ and $S_{WI} = -0.36$, for income (I), total expenditure (C) and wealth-adjusted income (WI), respectively. The graph in Figure 4 contrasts the three Lorenz curves.

[FIGURE 4]

Thus, as expected, the gasoline tax turns out to be more regressive if ability to pay is measured by wealth-adjusted income than if the same is measured by income. The difference is substantial, as it represents a 24% increase in regressivity as measured by the Suits index. What is more, the difference is rather sizable, representing a 140% increase in regressivity, if the comparison is made with the outcome resulting from using total expenditure in the lifetime perspective.

[21] Tax payments are derived for each household by first dividing quarterly gasoline expenditure by the monthly gasoline price averaged over the corresponding three months. We use monthly US average tax-inclusive gasoline prices published by the US Energy Information Administration.

[22] Figure C2, in the Appendix, shows the median tax payment as a proportion of the alternative ability to pay measures, by decile.

4.3.2 The incidence of the tax across age groups

On average, households with a young or an elderly head of household consume less gasoline than those with a middle-aged head of household (see Figure C3, in the Appendix). At the same time, households of the latter type tend to exhibit greater ability to pay (Figure 2 above). However, we here examine how the incidence of the gasoline tax varies across age groups, depending on the measure of ability to pay alone.

[FIGURE 5]

Figure 5 shows the relative tax burdens across age groups, by measure of ability to pay. The disequalizing effect of using total expenditure instead of income turns out to be somewhat limited. By contrast, when using wealth-adjusted income instead of income, (on average) older age groups systematically bear lower burdens than younger ones. This means that, in relative terms, the burdens borne by older (younger) households are overestimated (underestimated) if wealth is not considered in measuring ability to pay.

5. Conclusions

The literature on the distributional incidence of gasoline taxes, as well as more generally of carbon pricing, ignores wealth as a dimension of economic welfare and, thus, as a component of ability to pay. With reference to the US federal gasoline tax, we show that this is an important omission, which results in a significant underestimation of both the regressivity of the tax and its inequity towards younger people. Taking wealth into account exacerbates the regressivity outcome because the distribution of wealth is much more concentrated than that of income, which is the standard measure of current ability to pay, and all the more so of total expenditure, often used as a proxy for lifetime ability to pay. Taking wealth into account also reveals that, in relative terms, younger people actually bear greater tax burdens than those resulting from using income or total expenditure as measures of ability to pay. This is the case because, on average, older people own more wealth.

Our analysis is relevant to developed economies both with patterns of energy consumption across income distribution and wealth concentrations comparable to those in the US. The findings appear particularly important in light of the opposition to gasoline tax increases, or to the introduction or deepening of carbon pricing, motivated by the inequity of energy price increases. To be sure, to overcome this type of opposition, the distributional effects in question first need to be properly assessed. It will then be possible to address them through better calibrated redistributive measures. The relevance of our findings is further reinforced by the fact that a significant part of the literature draws conclusions pointing right in the opposite direction. Notably, the lifetime perspective taken by many empirical studies results in somewhat mitigated distributional effects, including, e.g., gasoline taxes turning from regressive to proportional. However, the utility of the policy implications that this type of result bears is questionable on different levels. First, the lifetime perspective is not well-suited for assessing the fairness of price changes in that people make interpersonal welfare comparisons – and hence equity judgments – based on realised outcomes, not expectations. Welfare programs are indeed calibrated based on observed welfare differentials, not expected ones. Second, the strong assumptions underlying the lifetime approach affect the accuracy of the outcomes.

Our analysis ultimately indicates that, by not considering wealth, the existing literature on the distributional incidence of gasoline taxes and of carbon pricing is biased against the regressivity of such policies. Greater regressivity than that emerging from this literature may actually help explain why equity concerns related to gasoline taxes and carbon pricing are such a big issue in the real world. Of course this does not make the cost-effectiveness case for environmental policies that raise energy prices any less powerful. It does imply, however, that their distributional effects should not be underestimated and that appropriate redistributive measures should be foreseen for the same policies to be fair and, therefore, ultimately for their implementation to be politically viable and sustainable.

References

Ando, A. and F. Modigliani (1963), “The life cycle hypothesis of saving: aggregate implications and tests”, The American Economic Review, 53(1), 55-84.

Arias, E. (2015), “United States life tables, 2011”, National vital statistics reports, 64(11), Hyattsville, MD: National Center for Health Statistics.

Aziparte, F. (2012) Measuring poverty using both income and wealth: a cross-country comparison between the U.S. and Spain. Review of Income and Wealth, 58(1), 24-50

Barker, T. and J. Köhler (1998), “Equity and ecotax reform in the EU: achieving a 10 per cent reduction in CO_2 emissions using excise duties”, Fiscal Studies, 19(4), 375-402.

Baumol, W.J. and W.E. Oates (1971), “The use of standards and prices for protection of the environment”, The Swedish Journal of Economics, 73(1), 42-54.

Bostic R., Gabriel, S. and G. Painter (2009), “Housing wealth, financial wealth, and consumption: new evidence from micro data”, Reg. Sci. Urban Econ., 39, 79–89

Bull, N., Hassett, K.A. and G.E. Metcalf (1994), “Who pays broad-based energy taxes? Computing lifetime and regional incidence”, The Energy Journal, 15(3), 145-164.

Bureau of Economic Analysis (2015), National Income and Product Accounts.

Callan, T., Lyons, S., Scott, S., Tol, R.S.J. and S.F. Verde (2009), “The distributional implications of a carbon tax in Ireland”, Energy Policy, 37(2), 407-412.

Chernick, H. and A. Reschovsky (1992), “Is the gasoline tax regressive?”, Institute for Research on Poverty Discussion Paper No. 980-92, University of Wisconsin-Madison.

Chernick, H. and A. Reschovsky (1997), “Who pays the gasoline tax?”, National Tax Journal, 50(2), 233-259.

Chernick, H. and A. Reschovsky (2000), “Yes! Consumption taxes are regressive”, Challenge, 43(5), 60-91.

Dissou, Y. and M.S. Siddiqui (2015), “Can carbon taxes be progressive?”, Energy Economics, 42, 88-100.

D'Orazio, M., DiZio, M. and M. Scanu (2006), “Statistical matching: theory and practice”, Wiley, New York.

Drews, S. and J.C.J.M. van den Bergh (2015), “What explains public support for climate policies? A review of empirical and experimental studies”, Climate Policy.

Ekins, P., Pollitt, H., Barton, J. and D. Blobel (2011), “The implications for households of environmental tax reform in Europe”, Ecological Economics, 70, 2472-2485.

Federal Reserve Statistical release (2012), Flow of Funds Accounts of the United States. Flows and Outstandings. Fourth Quarter 2011. Board of Governors of the Federal Reserve System, Washington DC.

Flues, F. and A. Thomas (2015), “The distributional effects of energy taxes”, OECD Taxation Working Papers, No.23, OECD Publishing.

Friedman, M. (1957), “A theory of the consumption function”, National Bureau of Economic Research, Princeton University Press.

Fullerton, D. and D.L. Rogers (1991), “Lifetime versus annual perspectives on tax incidence”, National Tax Journal, 44(3), 277-287.

Fullerton, D. and D.L. Rogers (1993), “Who bears the lifetime tax burden?”, Washington, DC: The Brookings Institution.

Hassett, K.A., Mathur, A. and G.E. Metcalf (2009), “The incidence of a U.S. carbon tax: A lifetime and regional analysis”, The Energy Journal, 30(2), 157-180.

Hassett, K.A., Mathur, A. and G.E. Metcalf (2012), “The consumer burden of a carbon tax on gasoline”, in Sterner, T. (ed.), “Fuel taxes and the poor”, RFF Press.

Kaplan, D. and A. Turner (2012), “Statistical Matching of PISA 2009 and TALIS 2008 Data in Iceland”, OECD Education Working Papers, No. 78, OECD Publishing.

Kennickell, A.B. (2007), “The role of over-sampling of the wealthy in the Survey of Consumer Finances”, in Irving Fisher Committee on Central Bank Statistics Bulletin 28, Bank for International Settlements.

Kosonen, K. (2014), “Regressivity in environmental taxation: myth or reality?”, in Milne, J. and M.S. Andersen (eds.), “Handbook of research on environmental taxation”, Edward Elgar.

Leulescu, A. and M. Agafitei (2013), “Statistical matching: a model based approach for data integration”, Eurostat Methodologies and Working papers, Publication Office of the European Union, Luxembourg.

Leuven, E. and B. Sianesi (2003), “PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing”.

Mathur, A. and A.C. Morris (2014), “Distributional effects of a carbon tax in broader U.S. fiscal reform”, Energy Policy, 66, 326-334.

Metcalf, R. (1999), “A distributional analysis of green tax reforms”, National Tax Journal, 52(4), 655-681.

Milanovic, B. (2002), “True world income distribution, 1988 and 1993: first calculation based on household surveys alone”, The Economic Journal, 112(476), 51-92.

Nordhaus, W.D. (1977a), “Strategies for the control of carbon dioxide”, Cowles Foundation Discussion Paper, mimeo.

Nordhaus, W.D. (1977b), “Economic growth and the climate: the carbon dioxide problem”, The American Economic Review, 67(1), 341-346.

Parry, I. (2015), “Carbon tax burdens on low-income households: A reason for delaying climate policy?”, in Clements, B., de Mooij, R., Gupta, S. and M. Keen (eds.), “Inequality and fiscal policy”, International Monetary Fund.

Piketty, T. (2014), “Capital in the twenty-first century”, Harvard University Press.

Poterba, J.M. (1989), “Lifetime incidence and the distributional burden of excise taxes”, The American Economic Review, 79(2), 325-330.

Poterba, J.M. (1991a), “Is the gasoline tax regressive?”, in Bradford D. (ed.), “Tax policy and the economy”, Vol.5, MIT Press, 145-164.

Poterba, J.M. (1991b), “Tax policy to combat global warming: on designing a carbon tax”, in Dornbush R. and J.M. Poterba (eds.), “Global warming economic policy responses”, MIT Press, 71-98.

Radner, D.B. and D.R. Vaughan (1987), “Wealth, income, and the economic status of aged households”, in: Wolff, E.N. (ed.), “International comparisons of the distribution of household wealth”, 93-120, Oxford University Press, New York.

Rässler, S. (2002), “Statistical Matching”, Springer, New York.

Rässler, S. (2004), “Data fusion: identification problems, validity, and multiple imputation”, Austrian Journal of Statistics, 33(1-2), 153-171.

Rosenbaum, P. R. and D.B. Rubin (1983) “The central role of the propensity score in observational studies for causal effects”, Biometrika, 70,41- 55.

Rausch, S., Metcalf, G.E. and J.M. Reilly (2011), “Distributional impacts of carbon pricing: A general equilibrium approach with micro-data for households”, Energy Economics, 33(S1), S20-S33.

Sterner, T. (2012a), “Fuel taxes, climate, and tax incidence”, in Sterner, T. (ed.), “Fuel taxes and the poor”, RFF Press.

Sterner, T. (2012b), “Distributional effects of taxing transport fuel”, Energy Policy, 41, 75-83.

Stiglitz, J.E., Sen, A. and J.P. Fitoussi (2009), “Report by the Commission on the Measurement of Economic Performance and Social Progress”, Commission on the Measurement of Economic Performance and Social Progress, Paris.

Suits, D.B. (1977), “Measurement of tax progressivity”, The American Economic Review, 67(4), 747-752.

Taussig, M.K. (1973), "Alternative measures of the distribution of economic welfare", Industrial Relations Section, Princeton University.

Tiezzi, S. (2005), "The welfare effects and the distributive impact of carbon taxation on Italian households", Energy Policy, 33, 1597-1612.

Verde, S.F. and M.G. Pazienza (2016), "Energy and climate hand-in-hand: Financing RES-E support with carbon revenues", Energy Policy, 88, 234-244.

Warren, A. (1980), "Would a consumption tax be fairer than an income tax?", The Yale Law Journal, 89(6), 1081-1124.

Weisbrod, B.A. and W.L. Hansen (1968), "An income-net worth approach to measuring economic welfare", The American Economic Review, 58(5), 1315-1329.

Williams, R.C., III (2016), "Environmental taxation", Discussion Paper DP-16-24, Resources for the Future, RFF.

Wolff, E.N. and A. Zacharias (2007), "The Levy Institute Measure of Economic Well-Being United States, 1989-2001", Eastern Economic Journal, 33(4), 443-470.

Wolff, E.N. and A. Zacharias (2009), "Household wealth and the measurement of economic well-being in the United States", The Journal of Economic Inequality, 7(2), 83-115

Wolff, E.N., A. Zacharias and A. Caner (2005), "Household wealth, public consumption and economic well-being in the United States", Cambridge Journal of Economics, 29, 1073-1090.

Wolfson, M.C. (1979), "Wealth and the distribution of income, Canada, 1969-1970", Review of Income and Wealth, 25(2), 129-140.

Zagorsky, J.L. (2005) Measuring Poverty Using Both Income and Wealth. Journal of Income Distribution, 3(4). 22-40

Appendix

A. Propensity score matching

Among the statistical matching methods, the literature distinguishes between parametrical and non-parametrical approaches, both with their advantages and trade-offs. We here apply a mixed method which takes the best of both worlds: the parsimony of parametric methods and the robustness to misspecification of non-parametric techniques (D'Orazio *et al.*, 2006). Specifically, we perform a propensity score matching (PSM) as described in Rässler (2002).

Though originally developed as a method to infer causal effects (Rosenbaum and Rubin, 1983), the PSM is increasingly being used to integrate pairs of datasets (Eurostat, 2013; Tedeschi and Pisano, 2013; Kaplan and Turner, 2012; Baldini *et al*., 2015). The PSM procedure consists of two steps. In the first step, a logit (or probit) model is fitted to a binary variable, *D*, indicating which of the two datasets an observation belongs to (e.g., D = 0 if observation *i* belongs to the donor dataset, D = 1 if *i* is from the recipient dataset). A set of selected variables, **X**, which are common to both datasets, are used as independent variables:

$$p_i = \Pr[D = 1 | X] = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)}} \tag{A1}$$

The propensity score (PS) is the predicted probability of an observation to belong to the recipient dataset – the CE sample in our case – conditional on **X**. The second step consists in matching the observations according to their PS, so that each unit of the recipient dataset is paired to the observation in the donor dataset exhibiting the closest propensity score according to a chosen distance function.

In our application, each CE observation is matched with one SCF observation. The wealth observed on the latter is then imputed to the former. The size of the SCF sample (30,075 observations) is much larger than that of our CE sample (2,179), which benefits the efficiency of the matching. However, due to oversampling of wealthy households in the SCF (Kennickell, 2007), we follow Bostic *et al.* (2009) in dropping the top income decile in the SCF dataset. This results in the removal of the top 5% wealth values (268 observations) and 3 observations with negative wealth.

A.1 Selection of the matching variables

The choice of the matching technique, such as the PSM, is only one of the steps required to integrate two datasets. The quality of the matching results is strongly dependent on the preliminary selection of the matching variables. These are a subset of the variables common to the two datasets (donor and recipient) selected based on both *a)* the similarity of their empirical distributions in the two datasets and *b)* the strength of their statistical association with the variables whose relationship is of interest, which are wealth and gasoline expenditure in our case.

To verify the first requirement, the common variables need to be homogenous across datasets both in terms of definition and statistical content. Thus, unless they are already identically defined, they have to be re-coded to ensure that the information they bear is exactly the same. Table A1 lists the common variables that have been considered as candidate matching variables.

[TABLE A1]

As the two datasets are samples drawn from the same population, the common variables should be homogenous in their statistical content too. That is, they should exhibit similar marginal and conditional distributions (Leulescu and Agafitei, 2013). Only variables with sufficiently similar distributions in the two datasets can be used in the matching algorithm. Different approaches can be used to assess the degree of similarity between pairs of distributions, the most popular being the simple inspection of the frequency distributions and the more rigorous calculation of the Hellinger distance. The Hellinger distance (HD) ranges between 0 and 1, these extreme values corresponding to perfect similarity and total discrepancy, respectively. In the literature, HD = 0.05 is often taken as reference threshold. Figure A1 shows the HD results obtained for the CE-SCF common variables, while also highlighting those eventually selected as matching variables (see below). Tables comparing marginal and conditional distributions of the same variables across datasets are available from the authors upon request.

[FIGURE A1]

The second requirement for a matching variable is to be statistically associated with both the variable of interest in the donor dataset, Y (wealth), and the variable of interest in the recipient dataset, Z (gasoline expenditure). After separately regressing Y and Z against the common variables, only those with sufficiently low HD (ideally below the 0.05 standard threshold) and, at the same time, showing significant explanatory power are selected as matching variables. To get plausible estimates of the (unobserved) joint distribution of Y and Z, strong explanatory power of the matching variables **X** is indeed critical[23]. In our application, the set of selected matching variables (i.e., those used in the logit model of the PSM) is narrowed down to: household income, housing tenure, age of the reference person, her education level, her marital status and her employer type.

[23] If the matching variables have strong statistical association both with Y and Z, the fundamental assumption of conditional independence between Y and Z (conditional on **X**) is easier to hold. If so, inference concerning the actually unobserved association is valid (Rässler, 2004).

A.2 Matching results

Once the matching variables have been selected, different matching algorithms can be considered. The choice of the matching algorithm is based on the quality of the resulting matching. This is usually assessed by three increasingly demanding criteria concerning the similarity between the distribution of the observed target variable (here, wealth) in the donor dataset and that of its imputed counterpart in the fused dataset. With reference to the target variable and the matching variables, the three criteria concern the preservation of i) the marginal and conditional distributions, ii) the correlation structure, and iii) the joint distribution (Rässler, 2004). In our application, while different matching methods perform similarly in terms of wealth's marginal and conditional distributions, the Mahalanobis metric (Leuven and Sianesi, 2003) outperform with respect to the other two more stringent criteria. Overall, the quality of the matching is deemed satisfactory.

(i) Marginal and conditional distributions

The HD between observed wealth in the SCF dataset and imputed wealth in the fused dataset is equal to 0.05, indicating only a small discrepancy. Table A2 contrasts the respective marginal distributions. The similarity between the two distributions is also illustrated with the Q-Q plot and the histogram in Figure A2 (top and bottom graph, respectively).

[TABLE A2]

[FIGURE A2]

Figure A3 contrasts the conditional distributions of observed wealth and imputed wealth in the donor dataset and in the fused dataset, respectively, against some matching variables. The conditional distributions in the first dataset are generally well preserved in the second.

[FIGURE A3]

Moreover, Table A3 contrasts the composition of observed wealth and that of imputed wealth. Again, the main distributional characteristics are maintained after the matching. Thus, the relative importance of the different wealth components (as well as the respective ownership rates) is similar in the donor and in the fused dataset.

[TABLE A3]

(ii) Correlation structure

The second, more demanding assessment criterion concerns the preservation, after the matching, of the correlation structure of wealth and the matching variables. Accordingly, Table A4 contrasts the relevant correlation matrices in the donor dataset (observed wealth) and in the fused dataset (imputed wealth). No major differences are observed.

[TABLE A4]

(iii) Joint distribution

Finally, the similarity of the joint distributions of wealth and the matching variables is assessed by regressing observed wealth and imputed wealth, in the respective datasets, against the matching variables. The statistical significance of the difference between the two sets of coefficients is then evaluated by means of a Hausman test. Table A5 shows the estimated coefficients of the two wealth functions as well as the outcome of the Hausman test. With reference to the latter, both at the .01 and .05 significance level, we fail to reject the null hypothesis that the coefficients are not systematically different. This result further validates the reliability of the matching performed.

[TABLE A5]

B. Rates of return on different assets

[TABLE B1]

[TABLE B2]

C. More results

[FIGURE C1]

[FIGURE C2]

[FIGURE C3]

FIGURES and TABLES

Table 1 – Net worth (per adult equivalent) and its composition.

	Mean	Std. Dev.	Min	Max	Mean share of Net worth	Ownership rates[c]
Net worth	217,293	413,973	-242,446	3,446,505	100%	100%
Assets						
Asset1: Houses[a]	108,533	151,996	0	2,500,000	50%	70%
Asset2: Other real estate and business[b]	49,184	203,522	-78,000	2,894,000	23%	28%
Asset3: Liquid assets	23,510	59,821	0	815,000	11%	94%
Asset4: Financial assets	27,502	113,193	0	1,764,000	13%	33%
Asset5: Retirement assets	54,562	149,937	0	2,123,001	25%	50%
Debts						
Debt1: Mortgage debt	37,897	69,116	0	890,666	17%	44%
Debt2: Other debt	8,103	19,928	0	450,000	4%	61%

a.: Houses refer to primary residences only. **b**: Other real estate consists of secondary residences, land, and rental property. Business refers to net equity in unincorporated business (both farm and non-farm). **c**: Percentage of households owning the asset.

Table 2 – Average composition of wealth-adjusted income (WI).

	Home owners (N = 1553)			Home renters (N = 626)		
	Mean	Std. Dev.	CV	Mean	Std. Dev.	CV
Money income (MI) of which Property income (PI)	73.1% (1.3%)	21.4 (8.3)	0.29 (6.13)	94.5% (1.1%)	18.7 (12.9)	0.19 (10.93)
Wealth annuity (WA)	16.1%	18.8	1.17	6.6%	14.2	2.15
Imputed rental income (IRI)	12.1%	10.5	0.87	NA	NA	NA

Table 3 – Distributions of Net worth and ability to pay measures (per adult equivalent).

	Net worth	Income	Total expenditure	Wealth-adj. income
1st percentile	-45,740	1,629	6,538	3,016
5th percentile	-14,497	7,064	9,739	8,382
10th percentile	-5,252	10,000	11,636	11,764
25th percentile	967	17,680	17,045	22,314
50th percentile	57,666	31,481	26,416	42,110
75th percentile	241,866	53,199	39,569	77,375
90th percentile	617,200	85,833	56,644	134,699
95th percentile	1,035,670	110,528	71,688	192,688
99th percentile	2,033,803	213,486	111,197	320,233
Mean	217,293	42,392	31,613	61,356
Std. Dev.	413,937	39,594	21,167	62,166
CV	1.91	0.93	0.67	1.01
Kurtosis	18.66	14.16	11.27	11.31

Skewness	3.50	2.74	2.15	2.49
Gini coefficient	0.76	0.44	0.34	0.47

Figure 1 – Lorenz curves of ability-to-pay measures.

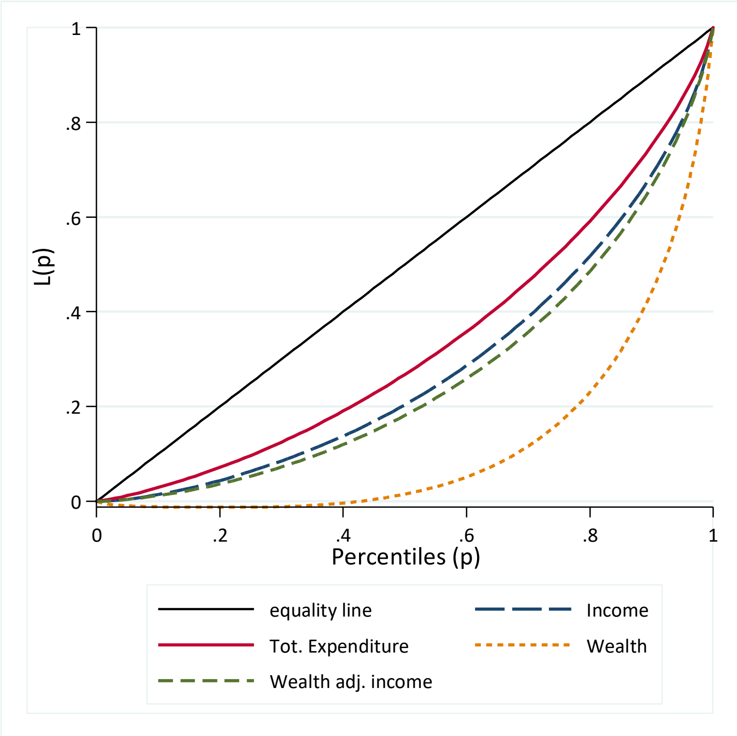


Table 4 – Frequency of changes in quintile ranking (%).

Change in quintile	Income vs Tot. exp.	Tot. exp. vs Wealth-adj. income	Income vs Wealth-adj. income	Tot. exp. vs Net worth	Income vs Net worth
-4	0.5	0.4	0.5	0.9	1.1
-3	2.0	1.7	0.8	4.1	3.7
-2	5.6	5.2	2.1	10.9	9.7
-1	16.7	20.3	10.7	20.3	20.0
0	46.3	45.2	67.2	30.6	32.2
1	23.0	19.6	18.6	17.3	18.9
2	5.2	5.7	0.2	9.7	8.4
3	0.6	1.6	0.1	4.1	4.3
4	0.1	0.4	0.0	2.2	1.6
Total	100.0	100.0	100.0	100.0	100.0
Correlation	0.72	0.66	0.88	0.34	0.42

Table 5 – Frequency distribution of households by head of household's age.

Age group	< 25	25-34	35-44	45-54	55-64	65-74	> 74
Frequency (%)	3.5	13.0	17.2	20.7	20.0	14.4	11.3

Figure 2 – Wealth and ability to pay measures by head of household's age group.

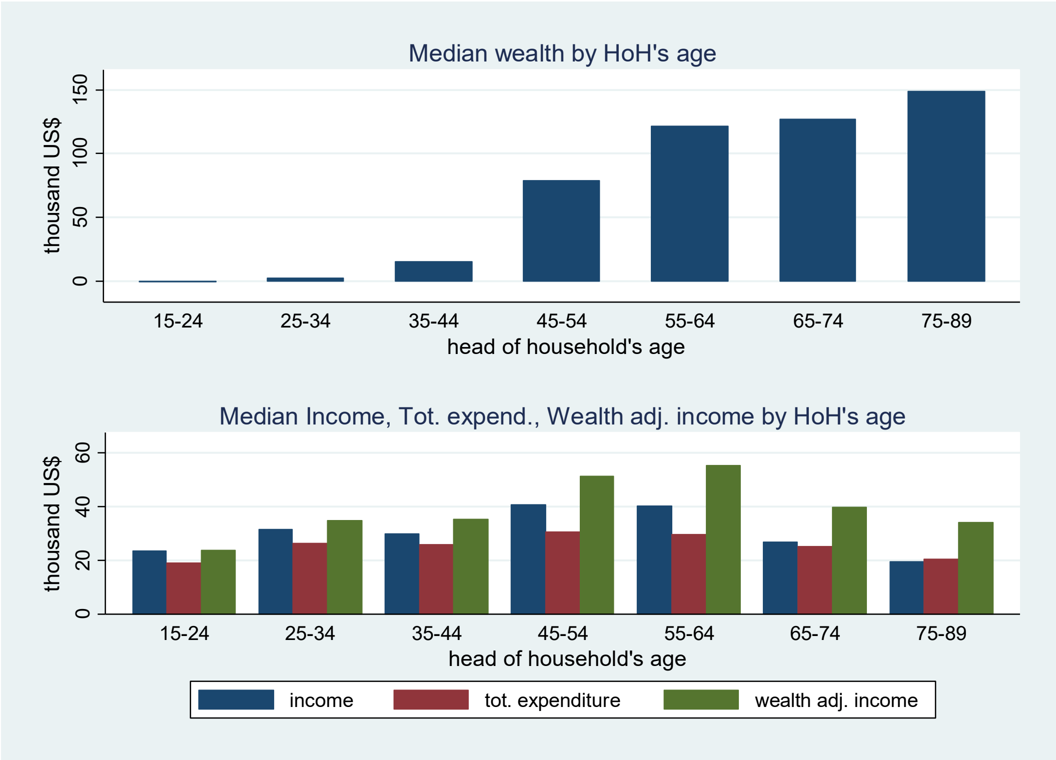


Figure 3 – Gasoline expenditure as a share of alternative ability to pay measures.

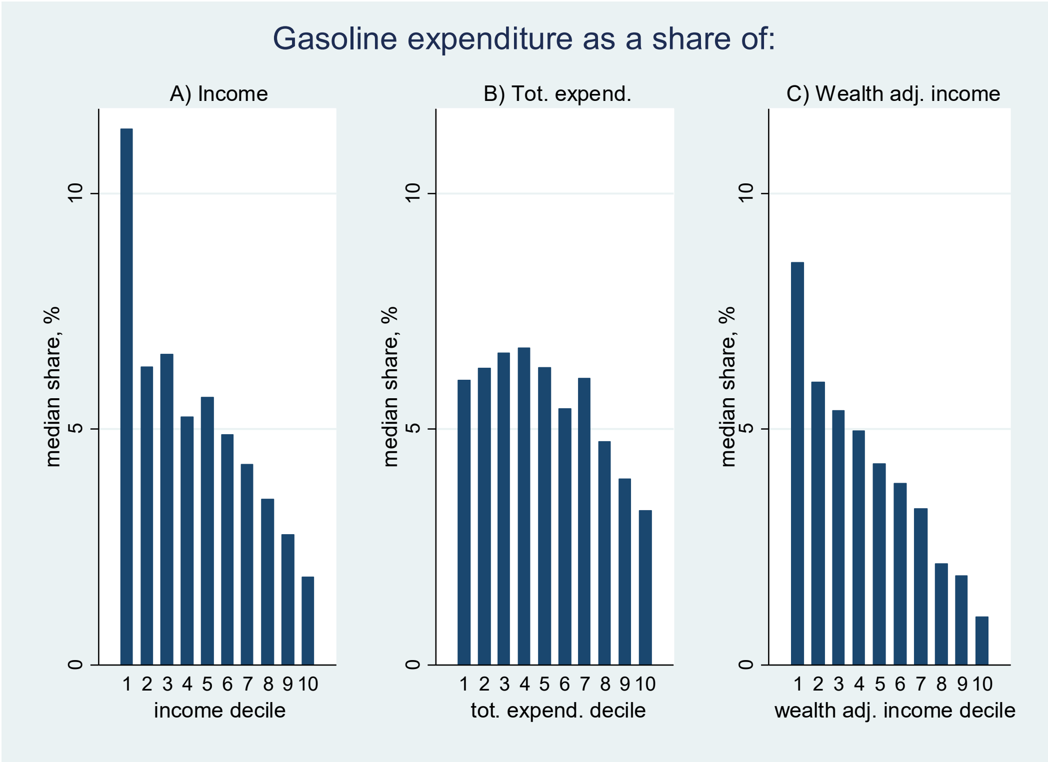


Figure 4 – Lorenz curve for the US gasoline tax, by ability to pay measure.

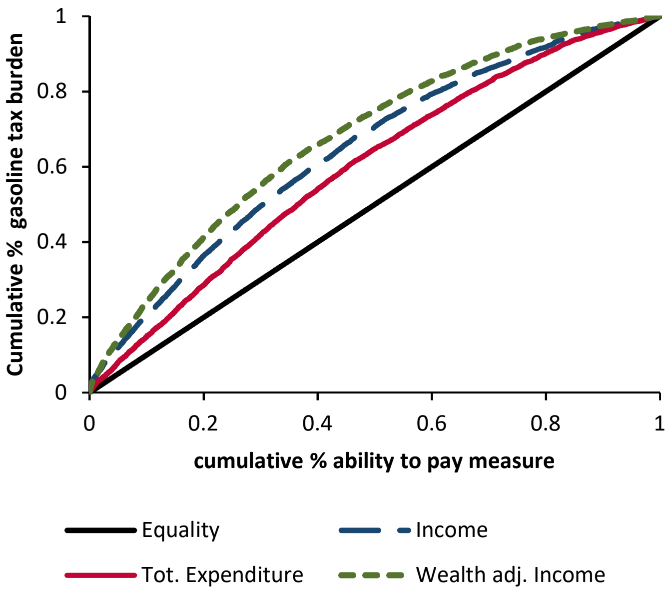
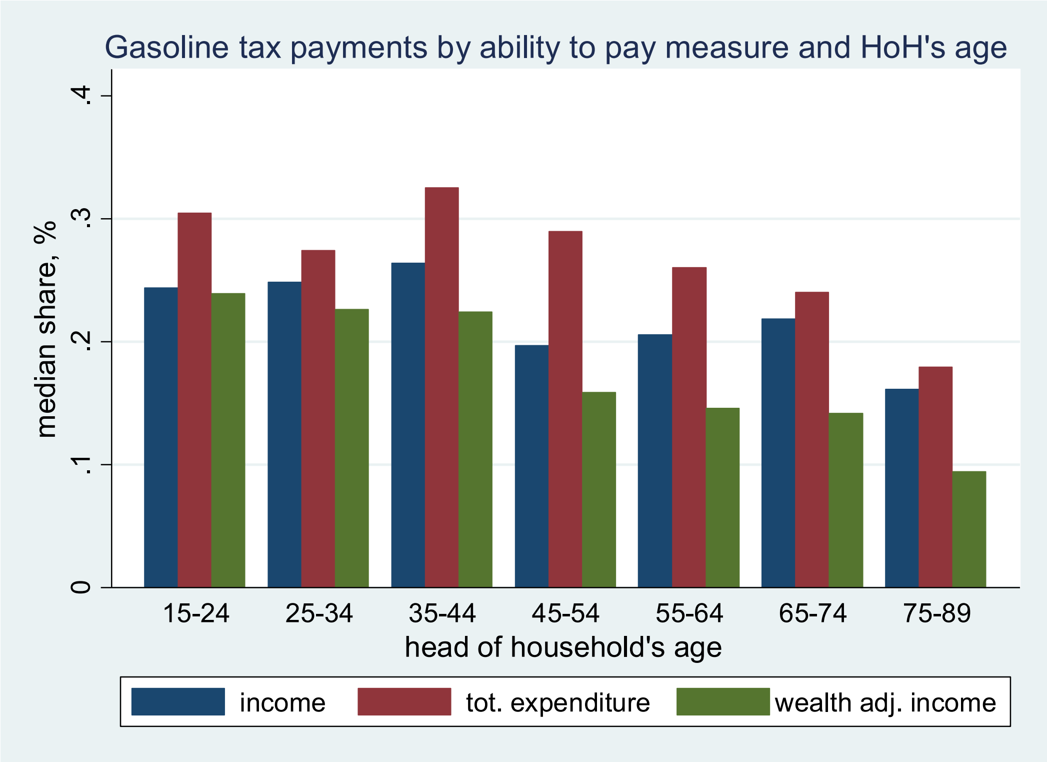


Figure 5 – Tax burdens as shares of alternative ability to pay measures, by head of household's age.



APPENDIX

Table A1: Common variables CE-SCF considered for statistical matching.

Characteristics of Reference person and Spouse	Marital status, Sex reference person, Sex of spouse, age of reference person, age of spouse, race of the Reference Person, race of spouse, education of Ref. Person, education of Spouse, ref. Person self-employed, Spouse self-employed.
Economic Characteristics	Family Income Before taxes, family Salaries, non-working group of ref. person (incnonw1_js), non-working group of spouse (incnonw2_js), hours worked in a week by ref., hours worked in a week by spouse, number autos, food at home, Food away.
House	Home renter (CU tenure), Rent paid
Household Structure	Family size, number of members under 18, number of members over 64.

Figure A1. Hellinger Distances for common variables

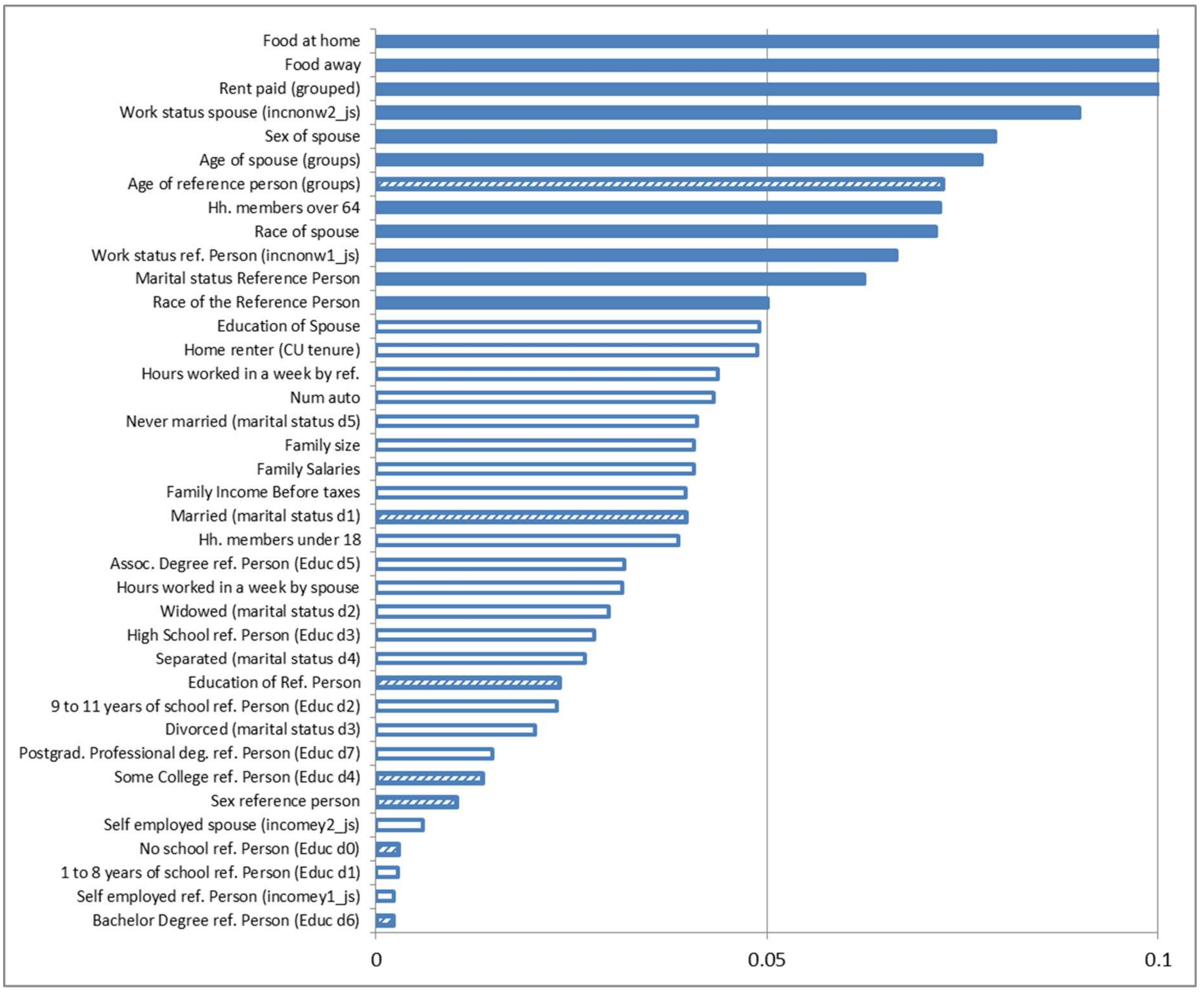


Table A2. Comparison of wealth distribution between CE and SCF after matching

wealth cutoff	SCF	Obs.	(%)	CE	Obs.	(%)	mean diff.	Diff/SCF(%)
-366360-	-58,418	1,209	5.10%	-59,147	74	3.40%	729	-1%
-18550-	-8,800	1,195	5.10%	-8,965	92	4.20%	166	-2%
-2420-	-406	1,184	5.00%	-313	102	4.70%	-93	23%
330-	1,979	1,203	5.10%	2,159	83	3.80%	-180	-9%
3810-	5,733	1,185	5.00%	5,480	95	4.40%	253	4%
8100-	10,276	1,199	5.10%	10,234	89	4.10%	42	0%
13260-	17,165	1,182	5.00%	17,174	97	4.50%	-9	0%
21800-	28,447	1,186	5.00%	28,109	111	5.10%	338	1%
35790-	45,522	1,164	4.90%	45,822	119	5.50%	-300	-1%
56900-	68,947	1,178	5.00%	69,256	111	5.10%	-309	0%
82110-	97,772	1,156	4.90%	98,056	129	5.90%	-284	0%
115140-	136,502	1,149	4.90%	135,499	135	6.20%	1,003	1%
159900-	185,634	1,162	4.90%	186,981	126	5.80%	-1,347	-1%

213600-	246,228	1,163	4.90%	248,734	120	5.50%	-2,507	-1%
282400-	331,172	1,166	5.00%	331,306	122	5.60%	-134	0%
387900-	450,216	1,167	5.00%	459,489	117	5.40%	-9,273	-2%
530800-	632,371	1,168	5.00%	629,761	120	5.50%	2,610	0%
735250-	886,546	1,169	5.00%	876,751	117	5.40%	9,794	1%
1063300-	1,361,926	1,163	4.90%	1,364,345	123	5.60%	-2,420	0%
1780000-	2,537,842	1,189	5.10%	2,596,898	97	4.50%	-59,056	-2%
Total		23,537	100%		2,179	100%		
Skewness	2.7			2.72			-0.02	-1%
Kurtosis	10.71			11.12			-0.41	-4%
Gini	0.75			0.72			0.03	4%
Theil (0)	1.52			1.33			0.19	13%
Theil (1)	0.91			0.84			0.07	8%

Figure A2. Wealth distribution in SCF and CE

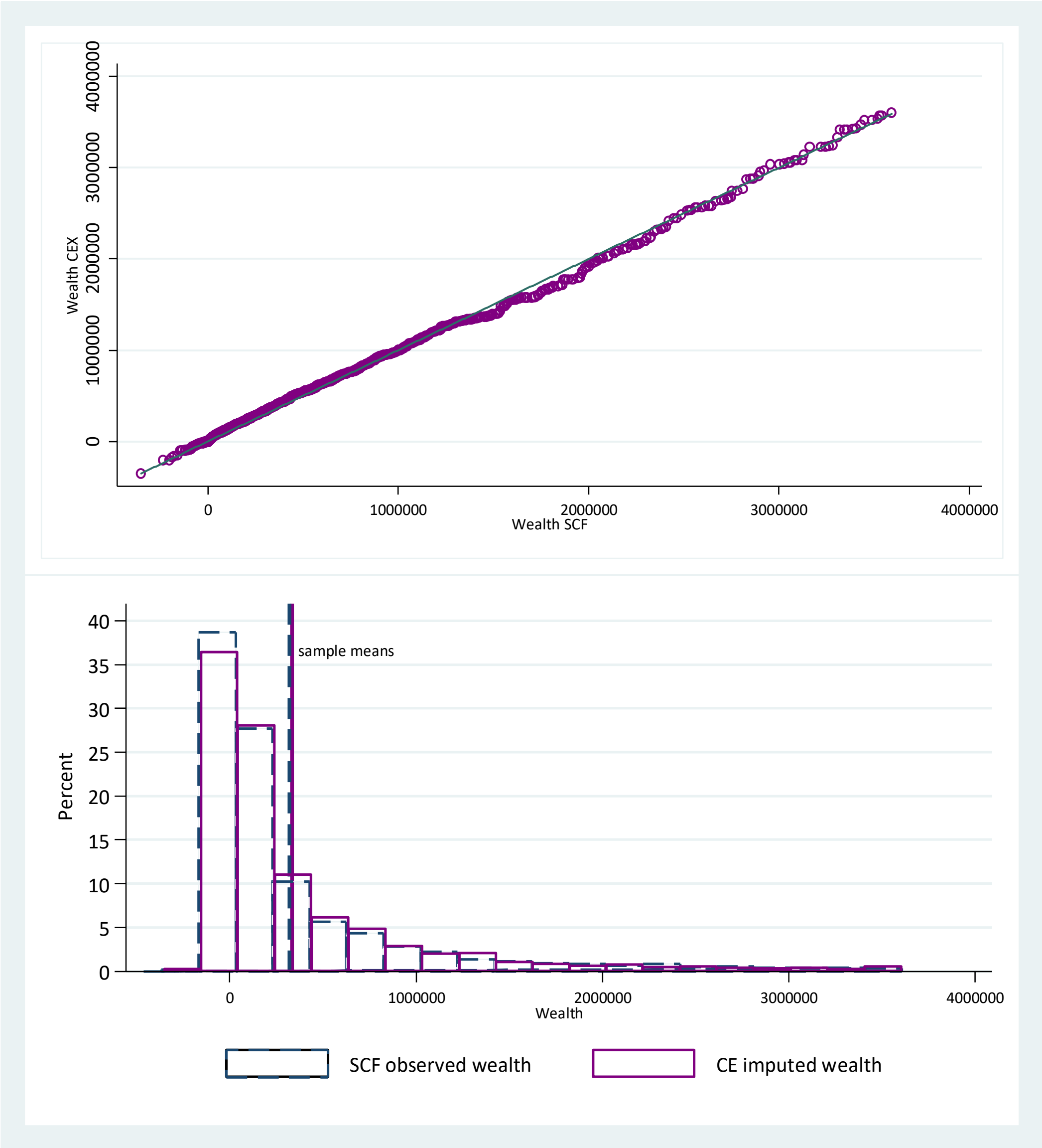


Figure A3. Conditional distribution of Wealth before and after fusion.

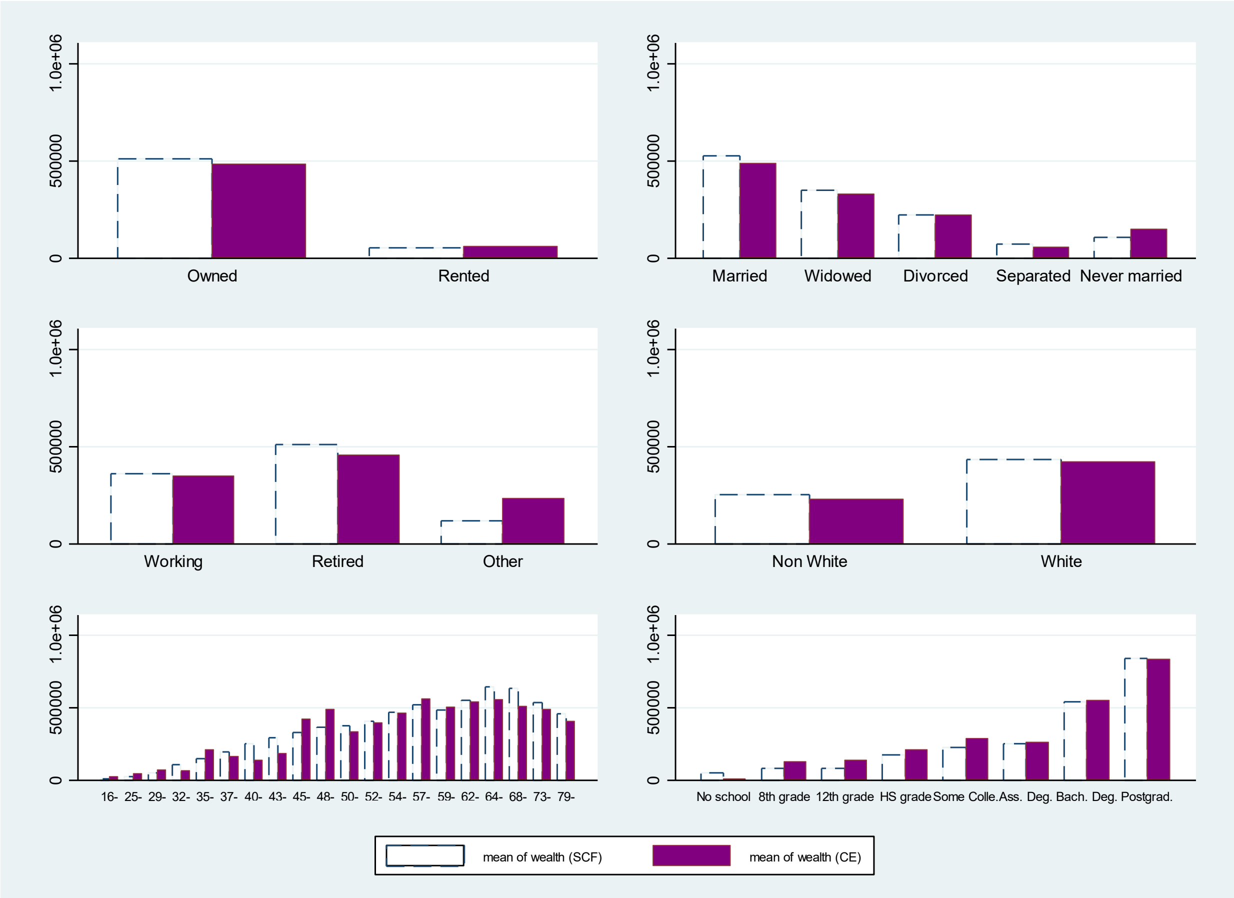


Table A3. Wealth components in imputed CE and observed SCF

Variable	Mean				Std. Dev.		Ownership rates	
	CE		SCF		CE	SCF	CE	SCF
Net Worth	217,293	100%	236,089	100%	413,974	473,921	100%	100%
Assets								
Owner-occup. house	108,533	50%	140,820	60%	151,996	200,514	68%	61%
Real estate and business	49,185	23%	50,803	22%	203,522	207,561	27%	30%
Liquid assets	23,511	11%	25,785	11%	59,821	72,965	93%	92%
Financial assets	27,502	13%	33,095	14%	113,194	147,581	33%	32%
Retirement assets	54,563	25%	61,781	26%	149,937	167,384	50%	48%
Debts								
Mortgage debt	- 37,897	-17%	- 61,870	-26%	69,117	11,398	42%	41%
Other debt	- 8,104	-4%	- 14,325	-6%	19,929	28,702	60%	63%

Note: values with sampling weights . **a.** Ownership rates refer to the percentage of households that actually own the given asset.

Table A4. Correlation structure between common variables in both SCF and CE after fusion

	Ln (netw.)	Ln (F. Income)	Sq. Income	Self. Empl. Ref	House tenure	Age ref.	sq. age ref.	No school	Some Coll	Bach. D.	Post.	marit.
ln(networth)	1.00											
ln(Family Income Before taxes)	0.59	1.00										
Squared Family Income Before taxes	0.35	0.58	1.00									
Self employed Ref.	-0.10	0.18	-0.01	1.00								
House tenure	-0.58	-0.32	-0.17	0.11	1.00							
Age Ref. Person	0.32	-0.02	-0.01	-0.46	-0.32	1.00						
Squared Age Ref. Person	0.29	-0.07	-0.04	-0.48	-0.29	0.98	1.00					
No school	-0.03	-0.05	-0.02	0.01	0.02	0.01	0.01	1.00				
Some College	-0.08	-0.10	-0.09	0.01	0.08	-0.05	-0.04	-0.02	1.00			
Bach. Degree	0.22	0.22	0.13	0.04	-0.10	-0.03	-0.04	-0.02	-0.23	1.00		
Postgrad.	0.29	0.33	0.26	0.04	-0.12	0.02	0.00	-0.02	-0.19	-0.21	1.00	
marital st. separated	-0.15	-0.13	-0.04	0.01	0.12	-0.05	-0.06	-0.01	0.03	-0.04	-0.06	1.00

	Ln (netw.)	Ln (F. Income)	Sq. Income	Self. Empl. Ref	House tenure	Age ref.	sq. age ref.	No school	Some Coll	Bach. D.	Post.	marit.
ln(networth)	1.00											
ln(Family Income Before taxes)	0.46	1.00										
Squared Family Income Before taxes	0.33	0.56	1.00									
Self employed Ref.	-0.03	0.36	0.14	1.00								
House tenure	-0.54	-0.24	-0.16	0.04	1.00							
Age Ref. Person	0.23	-0.20	-0.07	-0.54	-0.20	1.00						
Squared Age Ref. Person	0.20	-0.23	-0.10	-0.56	-0.17	0.99	1.00					
No school	-0.08	-0.05	-0.02	-0.04	0.04	0.04	0.04	1.00				
Some College	-0.03	-0.03	-0.06	0.02	0.07	-0.04	-0.04	-0.02	1.00			
Bach. Degree	0.19	0.21	0.16	0.13	-0.10	-0.10	-0.11	-0.03	-0.23	1.00		
Postgrad.	0.26	0.23	0.23	0.04	-0.07	0.04	0.03	-0.02	-0.18	-0.19	1.00	
marital st. separated	-0.12	-0.09	-0.04	0.02	0.06	-0.04	-0.04	-0.01	0.01	-0.04	-0.03	1.00

Table A5. Hausman test on wealth function between observed and fused wealth.

Dep. Variable: ln(Wealth)	(b)	(B)	(b-B)	
	Fused (CE)	Observed (SCF)	Difference	S.E.
Family Income Before taxes	0.000022	0.000021	0.000001	0.000001
Squared Family Income Before taxes	0.000000	0.000000	0.000000	0.000000
Self employed Ref.	-0.198223	-0.162168	-0.036055	0.081554
House tenure	-1.842897	-1.721890	-0.121007	0.081784
Age Ref. Person	0.016197	0.023399	-0.007203	0.011739
Squared Age Ref. Person	0.000106	0.000054	0.000052	0.000106
No school	-1.943236	-0.259389	-1.683847	0.615699
Some College	0.436587	0.380951	0.055636	0.089805
Bach. Degree	0.744419	0.828986	-0.084567	0.089676
Postgrad.	1.010688	0.901698	0.108991	0.107256
marital st. separated	-0.755710	-0.532961	-0.222750	0.250013
chi2(8) = (b-B)'[(V_b-V_B)^(-1)](b-B)	14.390			
Prob>chi2	0.072			

Notes: Ho: difference in coefficients not systematic.

Table B1 . Long term average rates of return (non-home wealth)

	Asset return rate *
Assets	
Real estate and business	2.54
Liquid assets	0.61
Financial assets	3.03
Retirement assets	2.79
Debts	
Mortgage debt	-3.81
Other debt	-3.81

Notes: Deflated values. These return rates consist in an updated version of those in Wolff and Zacharias (2009).

Table B2 – Comparison of home owners' imputed housing income (HI) to home owners' expected rent and home renters' paid rent (both reported in the CE).

	Expected rent	HI	Paid rent

Subsample	Home owners	Home owners	Home renters
Mean rent	10,660	7,503	6,431
Mean income	47,885	47,885	28,613
Mean rent/Mean income (ratio)	0.22	0.16	0.22
p10 (tenth percentile)	4,000	1,167	1,720
p25	6,000	2,767	3,024
p50	9,223	5,250	5,000
p75	12,800	9,933	8,544
p90	18,000	15,267	12,800
p95	24,000	21,250	16,800
N	1,519	1,553	626

Figure C1 – Changes in quintile ranking by head of household's age group.

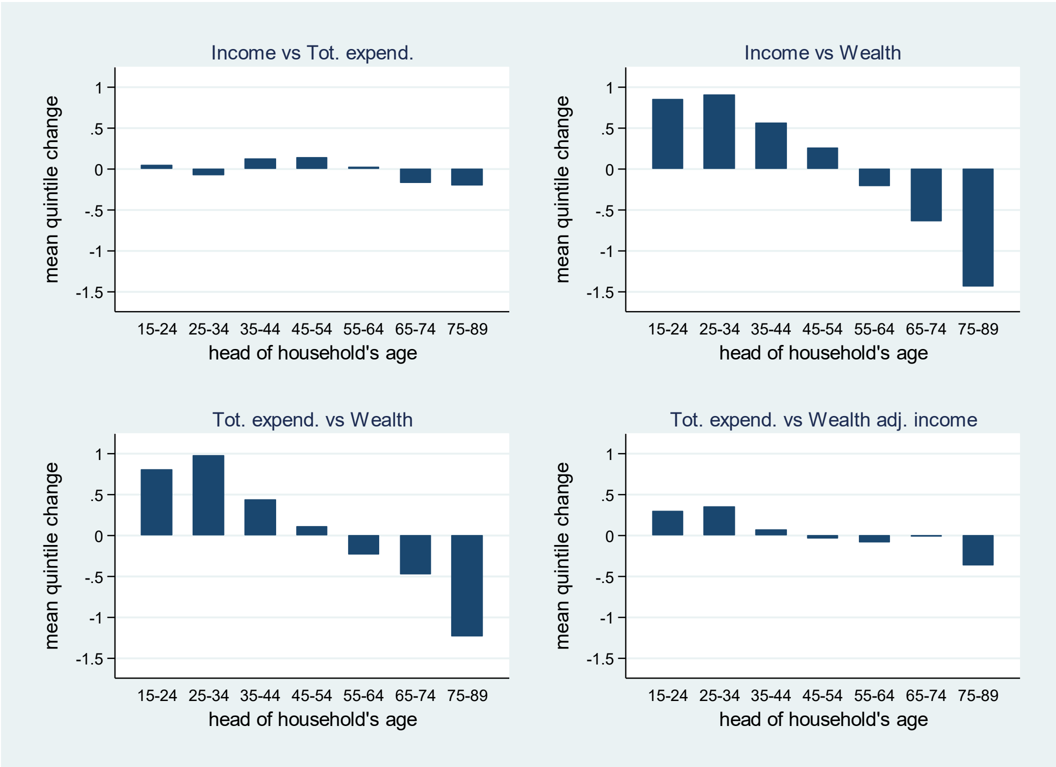


Figure C2 – Gasoline tax payments as a share of alternative ability to pay measures.

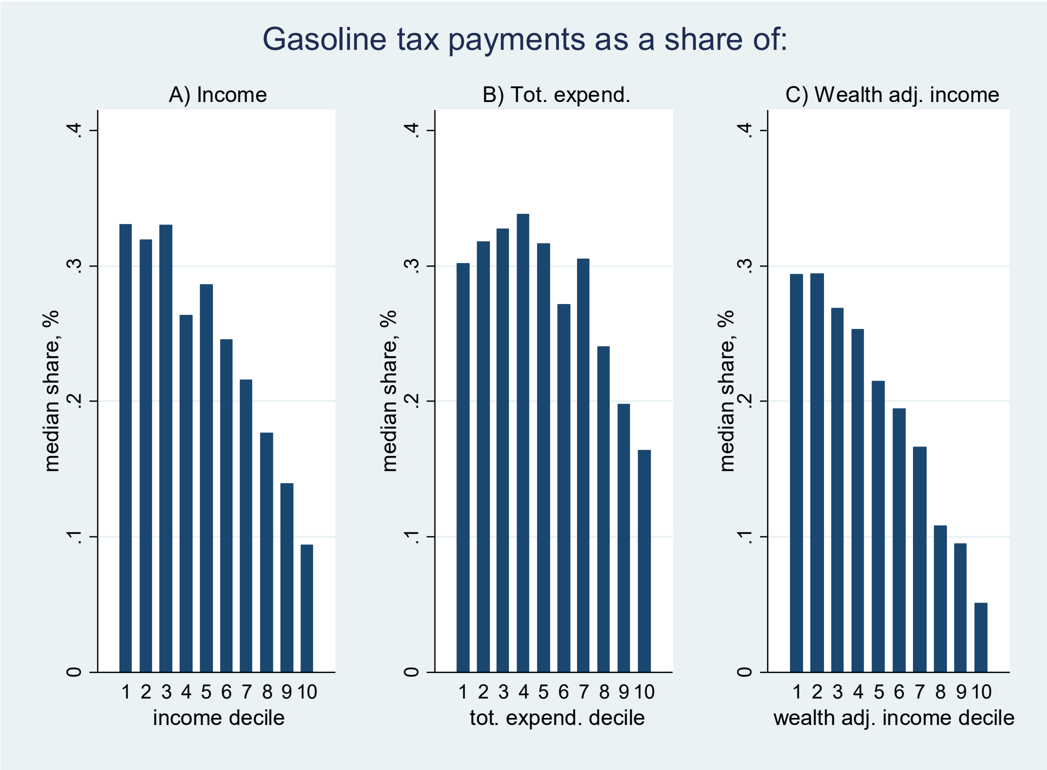


Figure C3 – Annual gasoline expenditure per adult equivalent by head of household's age group.