

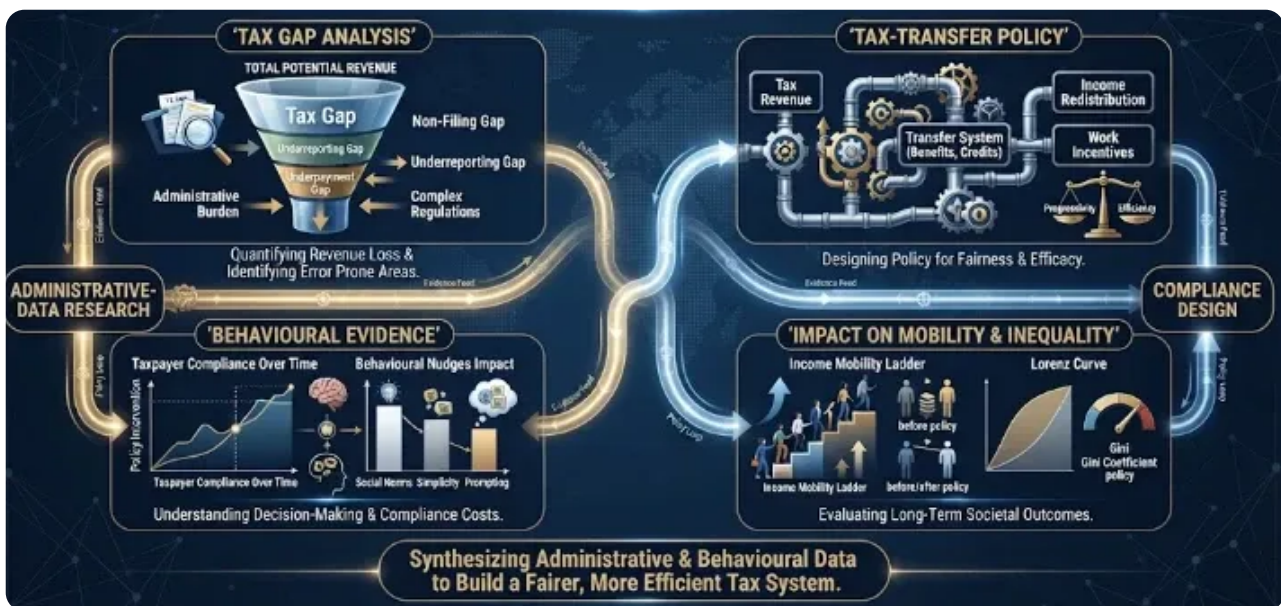
From Tax Gap to Compliance Design: Behavioural and Administrative Evidence on Tax-transfer policy, income mobility, inequality, administrative-data research

Kazi Rukaiya Sultana

Apr 12, 2026

Abstract

This article connects tax-gap analysis with inequality, income mobility, tax-transfer policy and administrative data. It argues that non-compliance has distributional consequences because the identity of those who underpay matters for the progressivity and legitimacy of the tax system. Deliberate evasion, aggressive avoidance, administrative error, cash-flow problems and exclusion from filing systems should therefore be distinguished rather than treated as one behaviour. Administrative data can support this diagnosis by linking information on income, businesses, benefits, employers and payments over time. Longitudinal records are especially useful because temporary low income differs from persistent disadvantage and can affect realistic debt or payment arrangements. The article also discusses evaluation of compliance interventions such as reminders, pre-filled returns, digital services and audits, with equity effects assessed alongside average outcomes. Data governance, privacy, algorithmic transparency and human review are highlighted as essential safeguards. Tax clinics provide complementary qualitative evidence explaining why a late return or debt occurred. Overall, the article promotes integrated evidence that connects revenue risk, household mobility, fairness, behavioural responses and administrative design.



Typically a tax gap is presented as a challenge in revenue administration, but the more policy impacts have far-reaching implications. Non-compliance in the long run shifts the tax burden, decreases transfer and public service funding, and potentially can affect competition between formal and informal activity. However, it also can be a burden on taxpayers with more limited means of income if data is missing or if there are complex rules to navigate. Thus a modern approach to compliance should take tax-gap issues and dimensions of inequality, income mobility, tax transfers and administrative data into a single policy approach.

Tax gap is the amount of difference between what tax should be collected according to current tax law and how much is actually received, typically broken down into gaps in gross tax and gaps in net tax. The methods of estimation are varied since there are different types of uncertainties, such as hidden income, inaccurate reporting, late payment, legal issues and complex corporate structures. A good estimate should, therefore, demonstrate methodology and limitations. The policy value is in understanding the patterns of risk – what taxes, sectors, transactions, taxpayer groups, etc., or what point in the compliance process, incur the highest losses.

Distributional analysis is a question that aggregate revenue can not answer. If those who don't comply tend to be the more affluent, of higher income and wealth or better able to use advanced tax planning, the actual tax system becomes less progressive than the law indicates. When the benefits of informalisation are more prevalent among lower income groups due to high costs of formalisation, a different approach to the policy is necessary. This may signal vastly different equity concerns, then, with the same amount of uncollected tax. There should be a clear difference between deliberate evasion, aggressive

avoidance, the administrative error, cash-flow difficulty and exclusion from available filing systems.

That is more possible with administrative data. Tax records include details of income, business income, withholding, deductions, payments, and filings. Social-protection systems will capture information on benefits and household conditions, and employer and financial reporting can provide third-party data. This has the potential to help with estimating tax-gaps, distributional analysis, programme evaluation and research into income mobility, provided it is linked securely and lawfully. The analytical opportunity is important because repeated records can provide an insight into changes in the same individuals or firms over time, not just one snapshot in time.

One reason why income mobility is important is because annual income inequality is only part of the picture. One household may lose their job, and then fall into low income for a brief period of time, or they might fall into low income and stay there for many years. The young worker can start at the bottom of the distribution and work their way up, or the other worker may experience repeated instability. Tax-transfer policy should differentiate between temporary shocks and disadvantage and persistent disadvantage. Longitudinal administrative data is better suited to identify those patterns than a one-year return.

The same information can be used to enhance work incentive assessment. Tax-benefit rules generate effective marginal rates which function as a decline in tax rates as income increases. Administrative data can document actual transitions into work, hours worked, benefit termination, and income growth, and models can estimate these rates. Researchers can assess results pre- and post-reforms, provided and within appropriate causal methods and privacy protections. Evidence can show whether there is a positive effect on employment from an in-work credit, whether there is an effect on employment from a benefit taper, or whether a benefit taper contributes to a low wage trap, or whether a filing requirement deters formalisation.

The same kinds of assessment methods can be used to evaluate compliance interventions. A reminder letter, a prefilled return, a third-party reporting rule, a digital service, an audit campaign, or a payment-plan reform can be applied to filing, payment and future compliance. Behavioural effects are important as the best intervention recovers more than one debt. It is able to alter future behavior at moderate administrative expense. Randomised trials can sometimes be used to assess communication strategies and quasi-experimental designs can be used to assess policy changes that cannot be randomised.

Equity should be included in those evaluations. Average compliance time can be reduced

and compliance may become more challenging for older individuals, individuals who have a disability, remote communities, and individuals whose primary language is not English, with a digital filing reform. While automated risk models can help to improve targeting, they may still lead to unfair results if the training data is based on historical enforcement practices. Administrative efficiency is therefore not enough. Error rates, outcome, access to service and burden across groups should be evaluated.

Another feedback loop is tax-transfer policy. More compliance will mean social spending or less compliance will mean lower rates, but the social spending choices is a political decision. Before any extra revenue is utilised, a compliance programme focused on high-income evasion could help to improve vertical equity. Conversely, an emphasis on easily spotted low-income mistakes may yield limited revenue and hurt trust. The expected yield and behaviour should be taken into account in risk based administration, but fairness and proportionality should be in strategic design.

One of the best examples of preventive compliance is where the third party reports. If information is reported separately by the employer, the bank, the platform or any other institution, the tax authorities can fill tax returns or detect mismatches. This type of system has a lower dependence on memory and decreases the chance for concealment.

Coverage still matters. Revenue sources not tracked well in third-party reporting may still be more difficult to measure, resulting in an uneven enforcement of regulations. But the size of data reporting should be expanded in light of material risk, privacy protection, and the compliance capacity of reporting entities.

As analytical capabilities grow, so does the need for data governance. Tax and Benefit records hold very sensitive data that can be used to build in-depth profiles of the households or businesses linked. Essential are legal authority, purpose limitation, access control, de-identification for research purposes, secure environments, retention rules and independent oversight. A technically good data set can be illegitimate if it is collected or connected without adequate protections. Trust in administrative data is like a valuable asset that can be squandered by mismanaged government.

This is also about transparency of algorithms. As analytics come into play and become a key tool for revenue agencies to prioritise cases, it's important to remember that risk-scoring should not be an unchallengeable black box. The taxpayer should be given clear and comprehensible grounds for the adverse action taken and the officials should have a way of monitoring the false positives. Even when a model indicates a complex case, human review is still crucial. It does not need to be disclosed how many anti-evasion

methods are used, but it does need to be disclosed how automated tools impact on enforcement.

Evidence that tax clinics and community services can add that is not captured in administrative data. A dataset documents a late return: a clinic can show that it was because of digital exclusion, family crisis, misunderstanding or that parents couldn't afford advice. A combination of quantitative and qualitative evidence should then be used. Administrative records reveal scale and patterns and frontline services provide mechanisms. That mix can make it difficult for policy makers to view each failure as an instance of the same behavioral issue.

The mobility analysis also facilitates improved debt policy. A household in temporary low income with good prospects for recovery would not be the same household as a household who were in permanent low income. Payment plans, hardship determinations, and collection practices can accommodate those differences without nullifying legal responsibilities. Predictive models should be used with caution but historical trends in income could help create more realistic pathways for administrators to design. It is possible to get more over the long haul from a sustainable arrangement than from aggressive action that leads to a viable small business going broke.

As of 12-04-2026, the road of tax gap to design for compliance more often leads through integrated evidence. Revenue estimates provide information on the magnitude of non-compliance, administrative data provide information on the distribution of risk and hardship, and longitudinal analysis indicates whether households rise or fall. The best system does that to make compliance easier, more effectively identify deliberate non-compliance, preserve the rights of the procedure and assess interventions following implementation. Addressing the tax gap isn't just about chasing more. It's about developing a revenue system that does that: delivers measurable fairness in the use of administrative powers, public value in the recovered resources, and effective burdens to meet legal intent.