# Referee Report on "The Distribution of Capital Gains in the United States"

## Summary

The paper studies the distribution of capital gains in the United States and their implications for inequality and tax progressivity. Using confidential IRS microdata from 2002-2021, as well as the U.S. Financial Accounts and the Survey of Consumer Finances, the authors construct comprehensive measures of household wealth and capital gains, including both realized and unrealized components across equities, housing, and private businesses. A key innovation in computing total capital gains is the development of new capitalization methods for heterogeneous valuation multiples and property tax rates across income groups, as well as a network-based approach to tracing partnership ownership structures.

The main findings are that capital gains are highly concentrated, with 75.7% accruing to the top 10% and 45.3% to the top 1% of individuals ranked according to Haig-Simons income – a measure of income that also includes capital gains in addition to factor income. Using the Haig-Simons income definition also raises the top 1% income share to 21%, compared to 18% without capital gains. Capital gains on equity, both public and private, are the main driver of the increased concentration.

There is also substantial heterogeneity in returns and capitalization rates, with richer individuals facing lower property tax rates, realizing higher returns on housing, and owning businesses that sell for higher multiples. The paper also documents that capital gains are lightly taxed. Because most gains remain unrealized or exempt, the effective tax rate is only around 5%. As a result Haig-Simons tax rates are essentially flat across the income distribution, undermining tax progressivity.

## Assessment

This is a very interesting paper on a topic of general interest that is at the center of many recent policy debates. The results, therefore, carry direct policy relevance, highlighting both the fiscal revenue potential of taxing capital gains more comprehensively and the distributional consequences of current exemptions and deferral. Of note is also the methodological innovation of estimating private business wealth using heterogeneous capitalization factors, which could also increase the paper's impact, although that methodology is already presented in a companion paper. Overall, I think the paper can make a nice and impactful contribution to the literature, and am therefore cautiously positive. That said, I have several major but concrete and, in my opinion, addressable concerns, which I would like to raise and which I think need serious consideration if the paper is to realize its ultimate potential.

## Comments

**1. Estimating public equity wealth and heterogeneous returns to equity.** The paper finds that capital gains on public and private equity – particularly on public equity, see Figure 6 (a) – are the

main drivers of the increased concentration of capital gains at the top of the income distribution. While the paper does a commendable job estimating private equity wealth it is much less ambitious when it comes to public equity wealth. I find this to be problematic in light of the finding that public equity capital gains are the most sizable and important. A related concern is that the authors assume that the returns to both public and private equity wealth are homogeneous. For private equity wealth, I find this assumption quite surprising given that the authors have information on the industry and size of the private business. It is well-known that different industries have differential comovements to the aggregate stock market (they have different market "betas"). A similar finding applies for firms of different size. In addition, firms of different size/industry load differentially on other factors that drive expected returns (e.g. the Fama-French High minus Low or Small minus Big factors). Therefore, estimating heterogeneous returns on private business wealth is quite feasible with the data the authors have, combined with standard data and results from empirical asset pricing. For public equity, there are recent contributions to the literature that allow for construction of heterogeneous capitalization factors of dividend income from tax returns, as well as heterogeneous expected returns on public equity – see the method developed in Chodorow-Reich, Nenov, and Simsek (2021 AER). These can be readily applied in the current context. I think they would matter for the concentration of capital gains, since it is well-established that older-richer individuals hold portfolios tilted towards higher dividend paying stocks. Such higher dividend paying "value" stocks also have a different comovement to the aggregate stock market compared to lower dividend paying "growth" stocks.

**2. The incidence of tenant-occupied residential real estate taxes.** Unlike owner-occupied real estate, for tenant-occupied real estate residential tax rates may be partly of fully pass ed through into rental rates. It is therefore unclear who bears the residential real estate tax on tenant-occupied housing. The pass through probably varies across housing markets. It may also impact the capitalization of tenant-occupied residential real estate. Can the authors try to incorporate this potential heterogeneity in tax incidence for tenant-occupied into their analysis or at least provide some robustness exercises that it is of second order? The paper on fiscal progressivity of U.S. federal and state governments by Fleck, Heathcote, Storesletten and Violante may be quite helpful in this regard.

**3. Use of tax return addresses for imputing owner-occupied real estate wealth.** The authors use the zip code address listed on the tax return to estimate the property value of owner-occupied housing. I have two concerns in this regard. First, some tax filers may not use their residential address as the tax return address and instead use a mail forwarding service. Second, some individuals may own multiple properties in different locations (e.g. vacation homes). The first concern is probably impossible to deal with but it would be good for the authors to at least discuss it as a potential limitation or possibly bound how problematic that could be. For the second, perhaps there is a way to bring some additional data to address this issue?

**4. Focusing the paper.** The paper shows many results and facts with a total of 26 figures. I think that's too much and it actually makes it harder to read the paper. Also some of the figures are not

easy to read given their size. I think the paper needs a bit more focus and a careful choice of what facts to present as figures, what to report as numbers in the text and what to move to the appendix (or not show at all).

**5. Interest-rate driven revaluations of equity and housing.** The paper argues that income changes from interest rate changes should not be counted towards income. For that reason the authors remove capital gains from fixed income assets from the analysis. However, interest rate changes may also drive the revaluation of equity and housing. Can the authors provide a bound on the possible effects from such revaluations given interest rate movements over their sample period? Much of the sample is during a very low interest rate period with small movements in interest rates so these effects are likely limited but I think they should be properly discussed.