How Should Commodities Be Taxed? A Counterargument to the Recommendation in the Mirrlees Review*

Spencer Bastani, Sören Blomquist and Jukka Pirttilä

Appendix 2 (Online appendix): Further computational results

In this section we investigate the robustness of the welfare gains of differentiated commodity taxation with respect to the number of types used to represent the wage distribution. For computational reasons we focus in this appendix on a model consisting purely of parents. This exercise can be seen as a first pass of examining the welfare enhancing effects of differential commodity taxes in the tagging optimum, i.e. those gains stemming from mitigating self-selection constraints within the parent-group in the tagging-optimum. The purpose of this appendix is not to offer fully fledged computational results; rather the idea is to scrutinize the effect of increasing the number of types. The numerical approach to setting up the optimal income tax problem is similar to Bastani (2013).

In the first half of Table A1 we show the effects of increasing the number of types, from 2, to 10, to 20, to 30, on the commodity tax structure and the associated welfare gains.[1] In all these calculations, just as in the calculations presented in the paper, we have used percentiles to approximate the wage distribution. In the exercise we present results both for the max-min case and the case with a Utilitarian social welfare function with harmonic weights (the weight on agent $i = 1, \dots, N$ is equal to $\rho^i = 1/i$).[2] As can be seen from the table, the welfare gains are lowest in the 2-type case, and then increase monotonically up to 1.98% in the weighted utilitarianism case and up to 2.74% when considering a max-min social welfare function.

Based on these results, one might argue that the reason the welfare gains change is because when using more types, the quality of the approximation of the underlying skill distribution improves. For instance, a 10-type approximation of the skill distribution (consisting of the deciles) will for

---

[1] The reason for not going beyond 30 types is computational and to ensure reliable results. Even in a model consisting purely of parents, we cannot rely on the single-crossing property.

[2] The reason for not focusing on a pure utilitarian social welfare function is that in such a case, none of the self-selection constraints would bind in the pure optimal tax solution. This might come as no surprise given that the utility function that we consider is quasi-linear in consumption goods. As is well known, the motive for redistribution is determined by the joint curvature of the utility function and the social welfare function.

obvious reasons capture the shape of the skill distribution better than a 2-type approximation consisting of the 33rd and 66th percentiles. However, it is also of interest, albeit for more technical reasons, to see how the welfare gains change with the number of types when we consider a series of approximations which capture the shape of the underlying skill distribution in an equally satisfying way. Our motivation for this exercise is that we want to investigate whether the welfare gains of differentiated commodity taxation obtained in a discrete model of optimal taxation are informative about the welfare gains that would obtain in a model with a continuum of types. For this reason, it is desirable to separate the effect on the welfare gains of the *number* of types from the effects on the welfare gains stemming from changes in the statistical quality of the approximation.

One way of measuring how well a discrete distribution approximates a continuous one is to compare the statistical moments of both distributions. When using a discrete approximation consisting of percentiles, it is the case that the moments of the approximation and the true distribution match better and better, as the number of points in the discrete distribution increases. To circumvent this issue, we have implemented a method to construct discrete wage distributions where we always match the first three moments of the underlying skill distribution, regardless of the number of types used in the approximation.[3] To understand the qualitative difference between this "matching moments"-procedure and the "method of percentiles" that we have used earlier we can note the following. When using percentiles, by definition, each discrete type is assigned a weight $1/N$ (where N is the total number of types). In contrast, a distribution obtained using a "matching moments" procedure typically contains a few points near the mode of the distribution with high probabilities assigned to them, as well as points in the far right tail of the distribution assigned with low probabilities. This latter procedure captures the mean, variance and skewness in a satisfactory fashion.[4]

In the bottom part of the table A1 we present the results of the simulations using this alternative way of representing the wage distribution, which also serves as an extra robustness check. As evident from the table, the welfare gains turn out to be very small for the 2-type distribution but

[3] Our method is based on Tanaka and Toda (2013). Our implementation is slightly different as we solve the approximation problem by finding the optimum to the primal optimization problem using constrained optimization techniques rather than solving the dual unconstrained problem. This turned out to produce more precise results.
[4] Note that when using a very large number of types a good fit is obtained regardless if the percentiles or the "matching moments"-procedure is used. When using a very small number of types it is not clear which method is "better", since with very few types we can only capture a small number of features of the wage distribution, and it is not clear which features are "essential" for the welfare gains of differential commodity taxation.

Table A1

Results using percentiles

# types	Weighted Utilitarian		Max-min	
	τ_2	Λ / GDP	τ_2	Λ / GDP
2	0.67	0.35%	0.96	0.93%
10	1.11	1.58%	1.29	2.22%
20	1.16	1.82%	1.33	2.41%
30	1.20	1.98%	1.28	2.74%

Results using a ”matching moments” procedure

# types	Weighted Utilitarian		Max-min	
	τ_2	Λ / GDP	τ_2	Λ / GDP
2	0.40	0.09%	0.50	0.18%
10	0.65	0.40%	0.96	1.20%
20	0.83	0.87%	1.19	1.99%
30	0.90	1.08%	1.27	2.30%

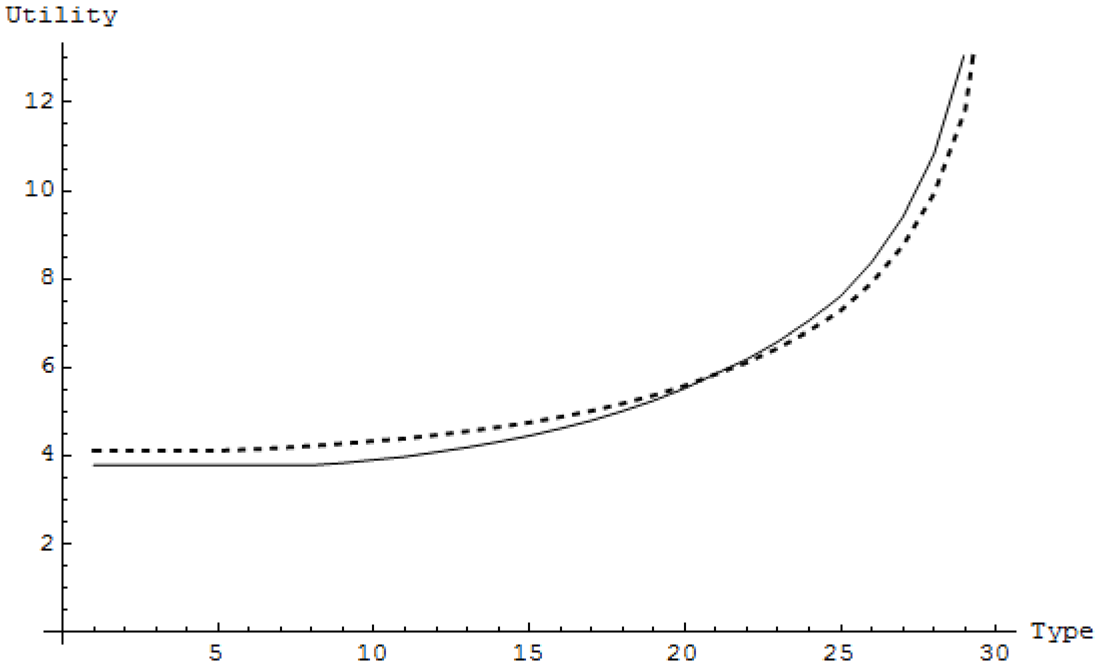
increase monotonically as the number of types increases to 10, 20, and 30. Our conclusion is that the welfare gains can be sizable also for this alternative way of approximating the wage distribution.

Source: own calculations.

The distribution of gains

In Figure 1 we show the winners and losers of a transition from the pure optimal tax benchmark with uniform commodity taxation to the differentiated commodity tax optimum under the restriction that child care is neither taxed nor subsidized. The results refer to the 30 type weighted Utilitarian case where the wage distribution were obtained using percentiles. The overall welfare gain as measured by the social welfare function is in this case equal to 1.98% of GDP. The figure shows the utility profiles for the pre-reform optimum (solid line) and the post-reform optimum (dashed line). Even though the differentiated commodity tax reform is not Pareto-improving, two thirds of all agents gain from the introduction of the reform in this particular model economy.

Figure 1



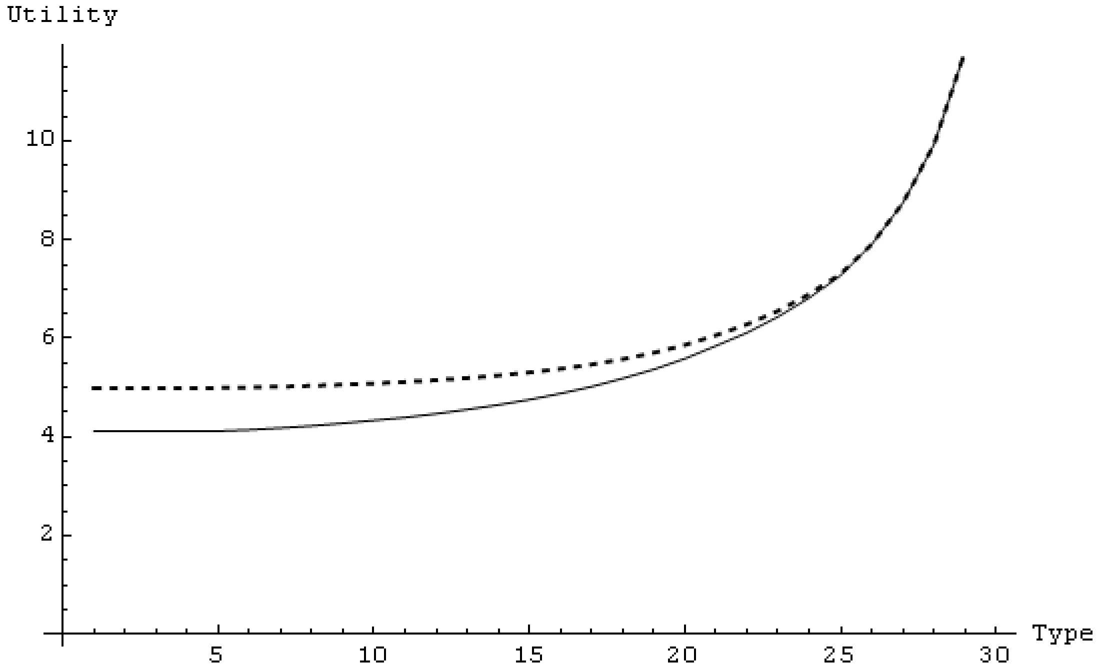
Source: own calculations

Pareto-improving reform

It should be noted however, that in this tagging scheme it is always possible to find Pareto improvements if we consider transitions from the pure optimal tax optimum (with uniform commodity taxation) to the "full" optimum (which as we have shown implies public provision and uniform commodity taxation). This is illustrated in Figure 2 below. The dashed line is the utility profile under the public provision optimum, and the solid line is the utility profile under the (pre-reform) pure optimal tax solution.[5] The utility in the public provision case is markedly higher for low and middle skill levels, and the move does not harm high skill level types.

[5] This solution was found by setting the subsidy rate on child care to 100% and maximizing the weighted Utilitarian social welfare function, imposing the constraint that each agent must be at least as well off as under the pure optimal tax solution.

Figure 2.



Source: own calculations.