Journal of Economic Behavior and Organization 220 (2024) 81–134

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Research Paper

Bearing the burden — Implications of tax reporting institutions on evasion and incidence ☆,☆☆

Kaisa Kotakorpi [a], Tuomas Nurminen [b], Topi Miettinen [c], Satu Metsälampi [d,*]

[a] Finnish Centre of Excellence in Tax Systems Research (FIT), Tampere University and CESifo, Finland
[b] Hanken School of Economics & Helsinki GSE and Finnish Institute for Health and Welfare, Finland
[c] Hanken School of Economics & Helsinki GSE, Finland
[d] University of Turku, Finland

ARTICLE INFO

JEL classification:
H21
H22
H26
D40
D44
D91

Keywords:
Tax evasion
Tax incidence
Third-party reporting
Double auction
Social image
Experiment

ABSTRACT

We investigate effects of tax reporting institutions on evasion and incidence using an experimental double auction market. We find that 28% of the sellers are truthful when only sellers report, but that 88% and 64% of them are truthful under costless and costly third-party reporting by buyers, respectively. Reporting behavior therefore responds to the intensity of deterrence. However, we find that prices do not fully reflect the lower taxes of the evaders. Thus, when sellers can unilaterally evade taxes, tax incidence deviates from the prediction of the standard model, and there is deadweight loss even if tax revenue is low. Pricing, incidence, and reporting patterns in all treatments are consistent with a model with heterogeneous lying costs, or a combination of lying costs and image concerns that give rise to a motivation to appear honest.

1. Introduction

Effective tax administration and enforcement are prerequisites for a well-functioning welfare state (e.g., Kleven, 2014). The so called tax systems approach to the analysis of taxation (Slemrod and Gillitzer, 2014), as well as literature on behavioral public finance more generally, emphasize that the reactions of economic agents to taxation may depend not only on tax rates and tax bases,

☆ We thank A. Brockmeyer, J. Harju, M. Hovi, P.H. Matthews, G. Van Moer, J. Slemrod, J. Sobel, and audiences at Helsinki GSE, MaTax, M-BEPS, MiddExLab Virtual Seminar, NCBEE, Stockholm School of Economics, VSE Prague, IIPF Congress 2019 and ZEW Public Finance 2021 for helpful discussions. We also thank P. Doerrenberg and D. Duncan for kindly sharing their z-Tree code with us and the PCRClab at the U of Turku for hospitality. Kotakorpi gratefully acknowledges financial support from the Research Council of Finland (grant nos. 277283 and 346250), Metsälampi from the Finnish Cultural Foundation (grant no. 00180719), Miettinen from the Yrjö Jahnsson Foundation (grant no. 20197200), Nurminen from the Emil Aaltonen Foundation (grant no. 190171), the Society of Swedish Literature in Finland (grant no. 1581), and the Yrjö Jahnsson Foundation (grant nos. 20177015 and 20197206).

☆☆ Earlier working paper versions of the paper have been circulated under the titles “The effect of reporting institutions on tax evasion: Evidence from the lab”, “The role of reporting institutions and image motivation in tax evasion and incidence” and “Bearing the burden — Implications of tax reporting institutions and image concerns on evasion and incidence”.

* Corresponding author.
E-mail addresses: kaisa.kotakorpi@tuni.fi (K. Kotakorpi), tuomas.nurminen@hanken.fi (T. Nurminen), topi.miettinen@hanken.fi (T. Miettinen), satu.metsalampi@utu.fi (S. Metsälampi).

https://doi.org/10.1016/j.jebo.2024.01.021
Received 9 January 2023; Received in revised form 17 December 2023; Accepted 22 January 2024

Available online 19 February 2024

but also on other design features of the tax system. Explaining these reactions may also require understanding how the intrinsic motivations of economic agents interact with the incentives generated by the tax system and other institutions (Coricelli et al., 2010; Dufwenberg and Nordblom, 2022; Attanasi et al., 2019; Bartling et al., 2022; Luttmer and Singhal, 2014; Bott et al., 2020).

In particular, third-party sources of income information such as records held by employers, business partners and credit card companies are considered critical for effective tax enforcement (e.g., Slemrod, 2007; Kleven et al., 2011). The importance of third-party information as a determinant of tax compliance has been acknowledged, yet literature utilizing randomized variation in reporting institutions is scarce. In addition, much of the empirical work on deterrence (e.g., Kleven et al. (2011) and subsequent field experiments reviewed in Pomeranz and Vila-Belda (2019) and Antinyan and Asatryan (2019)) focuses to a large extent on reporting responses. In general, reporting responses and real responses (i.e. responses pertaining to market behavior) are difficult to disentangle with field data.

We contribute to this literature by conducting a controlled laboratory experiment with randomized variation in taxation and tax reporting.[1] This provides valuable complementary insights to field studies, because reporting institutions can be randomized across subjects, and both market and reporting responses can be directly observed in the lab. Moreover, causal effects can typically be identified only for a short time period following (quasi-)experimental variation in the field. In the lab, one can compare behavior across institutions when market agents have frequent opportunity to learn about others' behavior and the implied audits and effective taxes. This permits studying causal effects when markets have settled to an equilibrium.

In more detail, we examine the impact of different tax reporting institutions on market clearing prices, quantities, tax incidence and compliance in the context of commodity taxation. We set up experimental induced cost-and-value double auction markets (Smith, 1962) in which a unit sales tax is levied on sellers. As benchmark conditions, we conduct a treatment without the tax and a treatment in which tax evasion is impossible. We then allow for tax evasion, conducting three treatments in which the initial tax payment is based on self-reporting by the seller, accompanied by a chance of an audit and possible fines.[2] In one condition the probability of the audit is exogenously set low. In the other two conditions, we allow also buyers to report their trades, and seller's audit probability increases if the seller under-reports relative to her customers. The audit probability is thus endogenously determined through the seller's own report and information provided by third parties (i.e. the buyers). A final treatment implements variation in the incentives to provide third-party information by imposing a cost on buyer reporting. This deviates from most of the previous literature which takes the existence (or not) of third-party information as given, and allows a study of whether reporting costs facilitate collusive tax evasion by buyers and sellers, and the associated consequences for market outcomes.[3]

Our design mimics some features of tax reporting institutions encountered in many countries (e.g. VAT-reporting), even though it is not built to exactly match the details of any given setting. We find a fairly generic approach preferable, cf. Section 3.1. The design builds on a well-established framework using a double auction market (Friedman, 1993) to analyze pricing behavior and tax incidence. Borck et al. (2002), Ruffle (2005), and Cox et al. (2018) provide evidence on tax incidence side equivalence in experimental competitive markets.

We find that when there is no buyer-reporting, many sellers evade at least some of their taxes due, but still prices are higher and quantities lower than in the case without taxes. When the reporting institution makes use of complimentary third-party information provided by buyers, we observe that both buyer and seller reporting are high and thus so is tax compliance, and market prices increase and quantities traded decrease. When reporting is costly to buyers, they report significantly fewer trades. However, somewhat surprisingly, tax compliance by sellers remains at a relatively high level.

We also find that the incidence of the effective tax is close to 50% as predicted by the standard model when taxes are automatically remitted and evasion is not possible. However, when sellers can unilaterally evade, the share of the tax burden borne by buyers is higher than in other treatments. This is due to the puzzling observation that even as sellers heavily under-report their trades, competition does not drive prices down to the Walrasian prediction. The unexpectedly high prices imply that there is a higher deadweight loss of taxation than the collected taxes would suggest. Neither markets nor the collection of tax revenue are efficient and effective. In the treatments where buyers also report their trades, tax compliance is high and the standard incidence result re-emerges.

A key finding of our study is therefore that unilateral tax evasion by sellers challenges the incidence result from standard tax theory, whereby incidence should be determined by demand and supply elasticities.

Our findings are not compatible with a standard self-interested and risk neutral model of decision-making. To discuss potential explanations for our findings, we first argue that they are not explained by risk aversion. We then draw on behavioral models that analyze the implications of lying costs and image concerns for decision-making. In particular, we apply a model that builds on Gneezy et al. (2018) and Abeler et al. (2019). The model analyzes lying behavior taking into account that even though agents may lie to increase their monetary payoff, lying entails an image cost as agents prefer being perceived as honest. Abeler et al. (2019) conclude that a model with both lying costs and image concerns is required to explain behavior in experimental literature that explicitly studies lying behavior (starting with Fischbacher and Föllmi-Heusi (2013)). While there is an extensive theoretical literature on lying behavior, models with image concerns have to our knowledge not been previously applied in the tax evasion literature, even

1 There is an extensive literature on laboratory experiments analyzing tax evasion; see e.g. Alm and Malézieux (2021) for a recent review.

2 The experimental taxes are paid out to the state tax authority to improve external validity.

3 Asking for a receipt may be associated with a small cost for signaling distrust, for instance, or credit card payment may involve a small price margin. The importance of buyer reporting costs is highlighted by Naritomi (2019) who finds in a Brazilian field setting that a small incentive to ask for a receipt increased the amount of third-party information and subsequent seller reporting. Our lab study allows us to further study market outcome and incidence implications of such changes.

though evasion is a prominent real-life example of dishonest behavior. We apply this type of a model to our double auction tax evasion setting, with separate market and reporting stages and show that it is consistent with our findings. We also argue that our results are consistent with a simpler model that incorporates lying costs only, if there is substantial heterogeneity in participants' lying costs across sessions. Lying costs affect market prices indirectly via their effect on the level of evasion and thus the amount of evasion-induced competition across markets.

Even though our design is not directly built to test the above models, and we cannot pin down exactly which version of the model provides the most accurate description of individual motivations in our setting, we find the theoretical analysis useful in providing some novel insights. First, we show that image concerns may have different implications for market vs. reporting behavior: even though the comparative statics of reporting behavior may appear completely standard in that individuals respond to deterrence, market responses may uncover intrinsic motivations that cannot be reconciled with a standard, self-interested model. The novelty here is that image costs are not suffered directly in association with tax reporting, which is not observed by other market participants. Image costs however may arise in the market stage, because low prices would convey a signal about tax evasion. This mechanism may explain why evasion is not in the end fully reflected in low prices, and gives rise to the non-standard incidence result.

Another novel insight, which is shared with the recent contributions of Tergiman and Villeval (2023), Halliday et al. (2021) and Benistant et al. (2022) is that when agents have heterogeneous lying costs, lower proneness to such intrinsic costs provides a competitive advantage in the market.

2. Related literature

Our work contributes to the literature on the determinants of tax evasion. In particular, we contribute to the study of the role of different reporting institutions in tax evasion, the debate concerning the importance of deterrence vs. non-pecuniary motivations behind compliance, and the interaction between evasion and market (or "real") behavior.

First, field studies on the role of third-party information in tax compliance mostly utilize naturally-occurring variation in third-party information stemming from certain types of incomes or transactions being subject to third-party reporting while others are not (Kleven et al., 2011; Pomeranz, 2015; Naritomi, 2019). Some recent studies have randomized the *salience* of third-party information in the field (e.g., Eerola et al., 2019; Harju et al., 2020). Alm et al. (2009) study individual income tax compliance in the laboratory in a setting where subjects differ in the fraction of their income that is subject to third-party reporting. All these studies find that third-party information increases compliance. We add to this literature by providing evidence on the deterrence effects of third-party information in a controlled lab experiment, where the institution that enables third-party reporting varies exogenously, while the actual supply of such information is determined by buyer incentives and behavior.

Our paper also relates to literature studying collusive tax evasion, which may undermine the effect of third-party information. Balafoutas et al. (2015) study how the efficiency in experimental credence goods markets is affected by trading partners' revealed intentions to evade taxes. Abraham et al. (2017) analyze the effect of social norms on joint tax evasion in the lab. In related field studies, Doerr and Necker (2021) conduct an experiment to study contractors' compliance and pricing behavior in online home improvement services markets, by implementing variation in the signals of consumers' willingness to collude. Bjørneby et al. (2021) implement a randomized audit study to provide evidence of joint tax evasion by workers and firms, while Paulus (2015) documents a similar phenomenon using survey and register data. In the present paper we vary the reporting institution (i.e. the cost of providing third-party information) to explicitly examine the conditions that may bring about collusion.

The papers on third-party information emphasize the role of deterrence in ensuring compliance. Whether the standard deterrence model is sufficient to explain tax evasion behavior is a long-standing debate in the literature. There is an extensive literature focusing on factors beyond deterrence, such as social and psychological factors (Luttmer and Singhal, 2014). Particularly relevant from the point of view of our discussion of image concerns are papers analyzing the effects of public shaming and public disclosure of tax information, including Bø et al. (2015), Perez-Truglia and Troiano (2018), Dwenger and Treber (2022), and Slemrod et al. (2022). In the field, public disclosure likely has effects that work both through social image and deterrence. Alm (2019) provides a review including a discussion of the role of social interaction both in theory and empirical literature on tax evasion.

Our paper also contributes to the literature on tax evasion and pricing behavior. Kopczuk et al. (2016) provide field evidence that differences in evasion opportunities between different sides of the market overturn the classical result that (nominal) tax incidence does not depend on the point of tax collection. Their study relates to the specific question of tax collection invariance, and the question studied in their paper is therefore different from ours. Their result arises if/when evasion has *any* kind of effect on prices, and market participants have different evasion opportunities. We examine the effect of evasion opportunities in general on effective tax incidence, and show that evasion may affect pricing behavior in ways that run counter to the predictions of standard theory. A paper close to ours is Doerrenberg and Duncan (2019) which analyzes the effect of evasion on incidence using a similar laboratory experiment. Like us, they find that providing sellers an opportunity to evade sales taxes implies that markets clear with higher quantities and lower prices than when evasion is not possible. They also find that the incidence of the effective tax falls more heavily on buyers when sellers have access to evasion.[4]

We add to these papers by providing lab evidence on the effects of tax reporting institutions on both evasion and market outcomes. In the lab, we can randomize institutions, and observe behavior after it has settled to an equilibrium. We confirm the importance

[4] In addition to the above mentioned empirical papers, Alm and Sennoga (2010) analyze the incidence effects of tax evasion using a computable general equilibrium model. Bayer and Cowell (2009) provide a theoretical analysis of the connection between enforcement policy and pricing behavior in an oligopolistic market.

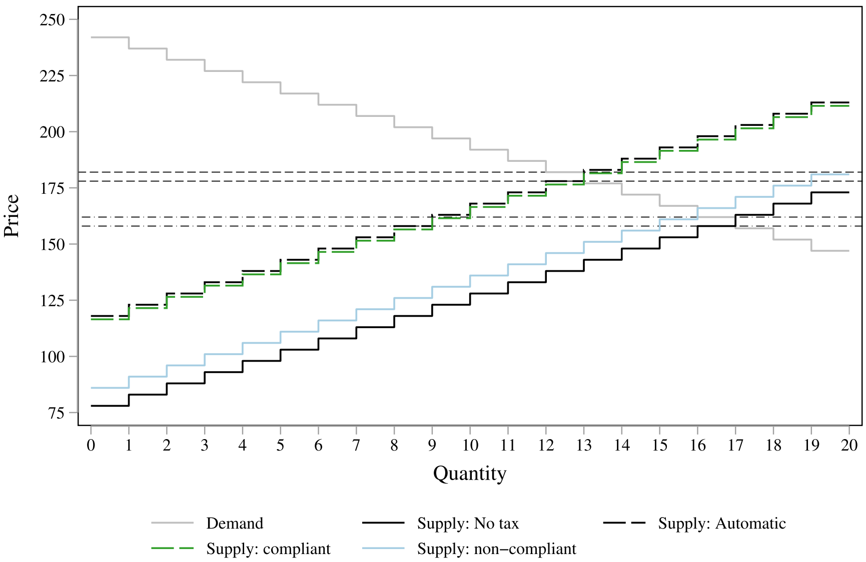

Fig. 1. Supply and demand. This figure plots the predicted market outcomes. Due to the discrete nature of the market, the supply and demand functions are step-functions, implying that the equilibrium prediction for the price is an interval. Dotted lower horizontal lines indicate the predicted price intervals in treatment No tax ([158,162]) and upper horizontal lines in treatment Automatic ([178,182]). Demand schedule (downward sloping light grey line) is fixed throughout the treatments. Solid black line depicts induced supply in the No tax and dashed black line in the Automatic treatment. Line above No tax supply depicts predicted supply in case of full non-compliance in Seller only and dashed green line overlapping supply in Automatic depicts predicted supply in case of full compliance in treatments Seller only, Seller + Buyer and Seller + BuyerC.

of deterrence through third-party information, emphasized in field studies, but find that the standard model is nevertheless not sufficient to explain market behavior in the presence of tax evasion.

Finally, we contribute to the behavioral literature on image concerns in lying (Fischbacher and Föllmi-Heusi, 2013; Gneezy et al., 2018; Abeler et al., 2019; Barron et al., 2022; Benistant et al., 2022; Halliday et al., 2021; Tergiman and Villeval, 2023). We apply a model of lying behavior with image concerns to tax evasion, and contribute by showing that this type of model can explain both the non-standard incidence result, as well as our other findings. We show that image concerns have different implications for tax reporting and for market behavior. Image concerns do not play a role at the reporting stage since reporting decisions are not observed by other market participants, while image concerns may affect pricing behavior. Tax evasion is thus reflected in a non-standard way in *market* behavior and market outcomes.

3. Experimental design

In each session, we conduct 25 repetitions of standard continuous laboratory double auction markets (Smith, 1962), with 5 sellers and 5 buyers trading units of a homogeneous good.[5] In a market each seller can sell up to 4 units and each buyer can buy up to 4 units of the good. Each unit k of seller i has a different per unit production cost c_{ik} and each unit k of buyer i has a different per unit reservation value v_{ik}. Cost and value schedules are randomly assigned to sellers and buyers, respectively, and vary across traders (see Table A.1 and Fig. 1 depicting the supply and demand curves).[6] The roles and cost/value schedules do not change during the experimental session and are private information.

The market opens for trading for 100 seconds in each period. Each buyer and seller can trade her units one at a time, starting with the first unit ($k = 1$), then the second ($k = 2$), and so forth. Traders may post offers to sell or buy at any time while the market is open. Each seller i may post an integer price $p_{ik} \in \{c_{ik}, ..., P^S - 1\}$ where P^S denotes the current standing ask (or 300 if there are no posted asks so far). Each buyer may post an integer price $p_{ik} \in \{P^B + 1, ..., v_{ik}\}$ where P^B is the current standing bid (or 0 if there are no posted bids so far). All traders observe the current standing bid and ask (if there are any).

A transaction takes place when a seller accepts a standing bid or a buyer accepts a standing ask. A seller can accept a current standing bid as long as $P^B \geq c_{ik}$. A buyer can accept a current standing ask as long as $P^S \leq v_{ik}$. When a transaction occurs, current standing prices are removed and the process of posting bids and asks begins again until the market closes. All accepted prices are displayed on traders' screens for the entire duration of the market stage.[7] Market participants are anonymous with respect to each other. This implies, in particular, that they cannot associate a given bid or ask with the participant who posted it. Communication is not allowed at any point.

[5] In addition, to help subjects get familiar with the environment and the interface, we ran three unpaid practice periods before the payoff relevant-periods.

[6] Costs and values are first randomized into sets of four costs and four values. These sets of four (see Table A.1) are then randomly assigned to traders at the beginning of each experimental session.

[7] See Appendix B for examples of trading screens.

3.1. Treatments

We have five treatments that differ in whether a per-unit sales tax is imposed on the sellers, and how the tax reporting institution is organized. The experiment was framed, i.e. we explicitly used the words "buyer", "seller" and "tax".[8]

In the No Tax treatment, there is no sales tax. A trader's market income in a given period is the sum of gross profits from each traded unit: $\Pi_i^S \equiv \sum_{k=1}^4 (d_{ik} p_{ik} - d_{ik} c_{ik})$ for sellers and $\Pi_i^B \equiv \sum_{k=1}^4 (d_{ik} v_{ik} - d_{ik} p_{ik})$ for buyers, where $d_{ik} = 1$ if the seller or buyer traded her k^{th} unit and $d_{ik} = 0$ otherwise. In the Automatic treatment, a per-unit sales tax τ, equal to 40 experimental currency units (ECU), is imposed on the sellers.[9] The tax is automatically collected, making tax evasion impossible. Hence a seller's market income is given by $\Pi_i^S \equiv \sum_{k=1}^4 (d_i p_{ik} - d_{ik} c_{ik}) - \tau s_i$, where $s_i \equiv \sum_{k=1}^4 d_{ik}$ denotes the number of units the seller sold in the current period.

In the Seller Only treatment, sellers are asked to file a tax report stating how many units they sold in the current period.[10]

A seller can report $r_i \in \{0, ..., s_i\}$, where r_i denotes the number of units seller i reports as sold in period. The sales tax is collected for each unit *reported* as sold, unless an audit is conducted, in which case the tax is collected for each unit actually sold. In addition, an audit implies the seller is fined $f = 40$ ECU for each sold unit the seller failed to report. The probability of an audit is exogenously fixed 10%, and it is independent across sellers. Now seller i's market income from trading in a given period is given by

$$\Pi_i^S = \begin{cases} \sum_{k=1}^4 (d_{ik} p_{ik} - d_{ik} c_{ik}) - \tau r_i, & \text{if seller } i \text{ is not audited} \\ \sum_{k=1}^4 (d_{ik} p_{ik} - d_{ik} c_{ik}) - \tau s_i - f(s_i - r_i), & \text{if seller } i \text{ is audited.} \end{cases}$$

The Seller + Buyer treatment is otherwise as the Seller only treatment but with an endogenously determined audit probability. More specifically, for each unit they bought in the current period, buyers are asked to decide whether to costlessly report or not report that unit.[11] The probability of an audit for a seller is 10%, unless the seller reports fewer sold units than the buyers who bought from her, in which case the audit probability is 80%.[12] The seller only learns whether she was audited, not what her trading partners reported.

Finally, the Seller + BuyerC treatment is otherwise as the Seller + Buyer treatment but reporting is now made costly to the buyers. If a buyer reports a positive number of units, she incurs a fixed reporting cost of 10 ECU. The cost corresponds to a hassle cost when filling out a tax report which, by revealed preference (Benzarti, 2020; Kotakorpi and Laamanen, 2017), may result in non-reporting. In this context, a fixed cost of reporting is very natural.

A few notes on our design are in order. First, we use the continuous double auction market institution, as it is well-known to generate (Walrasian) competitive equilibrium outcomes after convergence has taken place. This type of design therefore provides a well-founded setting to analyze the implications of tax incidence and has been employed in this context by Borck et al. (2002), Cox et al. (2018) and Doerrenberg and Duncan (2019).

Second, our design mimics some features of tax reporting institutions of many OECD countries, even though it is not built to exactly match the details of any given setting. For example, the Seller + Buyer treatment has similarities with a VAT-reporting scheme where also buyers report trades. In many countries, VAT-liable buyers are incentivized to report their trades by the possibility to deduct VAT payments from taxable profit. We do not explicitly incorporate this specific incentive structure, but we do implement variation in reporting incentives for buyers by comparing the effects of free vs. costly reporting. There are also some pros to adopting a more generic setting, in that an exact match between the design and some specific reporting context would limit its applicability to other settings. For example, similar seller-buyer reporting situations naturally arise in the labor market, where employee income is typically reported by both employers and employees. Even though third-party information is regarded as an effective deterrent on seller reporting in these types of settings, it does not fully preclude evasion, cf. Paulus (2015). In all treatments with taxes, to improve external validity, the participants are informed that the experimental tax revenue is paid out to the state tax authority.[13]

Third, even though participants are anonymous, the asks, bids and realized trading prices are observable to other market participants. This makes image motivation consistent with our framework. In earlier research, it has been found that (social) image motives may matter in experiments even under anonymity. For example, Fischbacher and Föllmi-Heusi (2013) find that people do not lie maximally even in a double-blind treatment in which anonymity is strictly preserved. This can be seen as evidence for unwillingness to be perceived as dishonest/opportunistic. Barron and Nurminen (2020) note that many subjects lie to their peers about their free riding in a public goods game despite being anonymous and not having monetary incentives to do so. Social image motivation may

---

8 See Appendix E for English translations of instructions for treatment Seller + BuyerC. The full set of instructions is available from the authors.

9 The tax equals 25% of the predicted median market equilibrium price of the No Tax condition and approximately 32% (21%) of the median cost (value). It therefore closely corresponds to the general VAT rate of 24% in Finland.

10 See Fig. B.3 in Appendix B for an example of a seller's reporting screen.

11 See Fig. B.4 in Appendix B for an example of a buyer's reporting screen. In particular, there is no option of saying "I prefer not to tell/report" which might be a way for buyers to achieve collusive tax evasion without having to lie.

12 We select the 80% audit probability mainly for two reasons. First, we want the audit probability triggered by third-party information to be sufficiently high to create enough contrast with the baseline 10% audit probability. Second, we choose a probability below 100% to reflect the fact that the audit may fail to detect the full extent of evasion. One can interpret the audit probability as a combination of a 100% audit probability and an 80% probability with which the tax authority detects any evasion. Moreover, we did not want to use 100% audit probability to avoid the certainty effect (Kahneman and Tversky, 1979).

13 The exact choice of terminology here involves a tradeoff between external validity on the one hand, and on the other hand providing the participants an accurate description of the experiment. The fact that the experimental "tax revenue" is indeed paid out to a government tax account increases external validity. On the other hand, we use the word "donate" in the experimental instructions to describe this transfer, to avoid impression that participants actually would have a legal liability to pay the tax.

Table 1
Main empirical hypotheses.

	No tax		Seller only		Seller + BuyerC		Seller + Buyer		Automatic
Predicted price	[158,162]	<		≤		<		<	[178,182]
Predicted quantity	17	>		≥		>		>	13

Notes: Summary of predicted market clearing prices and quantities in treatments. The predictions correspond to the main hypotheses stated in pre-registration (Kotakorpi et al., 2018), apart from treatment Seller + BuyerC which was not pre-registered (see also footnote 15).

be important in anonymous laboratory environments, for example, because 'rule rational' participants import into the lab behavior that is beneficial outside the lab (Grossman, 2015). In a setting where offers and trades are directly attributable to specific market parties, one would expect the effects of image motivation to be stronger (Casal and Mittone, 2016).

3.2. Procedures

We conducted 30 sessions with 10 subjects in the PCRClab of the University of Turku. In each session, we implemented one treatment condition. A total of 300 subjects, predominantly students at the University of Turku, participated in the experiment. There were 6 sessions and 60 participants in each treatment. Summary statistics of participants' demographic variables are shown in Table A.4. Participants were solicited through an online database using ORSEE (Greiner, 2015), and the experiment was run using the experiment software z-Tree (Fischbacher, 2007). After the experiment, subjects filled out a short questionnaire on background characteristics, attitudes towards cheating on taxes, trust and risk preferences.[14] Sessions lasted up to 110 minutes, and participants earned, on average, 10.00 EUR for the experiment, including a 5 EUR show-up fee. All tax revenue collected in the experiment was donated to the Finnish State Treasury, and this was common knowledge among the participants.

4. Predictions

In this section we present theoretical predictions concerning market prices, quantities and reporting behavior in our treatments. The theoretical predictions are derived under the assumption that the traders are money-maximizing and risk-neutral. This assumption allows us to derive precise predictions on reporting behavior in all treatments. However, since there is extensive evidence that not all participants in economic experiments are money-maximizing and risk-neutral, our pre-registered hypotheses acknowledge that there will likely be variation in the participants' preferences. The pre-registered hypotheses for the treatments with tax reporting are therefore not exact. Instead, the pre-registered predictions for a given treatment with reporting are expressed relative to other treatments (e.g., that market prices in the Seller only treatment are higher than in the No tax treatment).[15] The theoretical prediction coincides with pre-registered prediction in No tax and Automatic. All predictions reflect equilibrium behavior which markets converge to after experienced traders have had the opportunity to learn and adjust their responses to the behavior of other traders, audits and the implied effective tax rates. We derive the theoretical predictions formally in Appendix C.

Fig. 1 illustrates the predictions in our experimental markets (see also Table 1). For treatments No tax and Automatic, standard economic theory offers precise quantitative predictions. In the absence of the sales tax, only marginal cost matters for sellers' supply decisions. Market supply without tax (solid black line in Fig. 1) equals demand at 17 units in the Walrasian equilibrium, corresponding to a market clearing price between 158 ECU and 162 ECU. Imposing an automatically collected 40 ECU per-unit sales tax implies a 40 ECU upward shift in the market supply curve (dashed black line in Fig. 1). Therefore, in the Automatic treatment, market clears with a lower quantity of 13 units and a higher price between 178 ECU and 182 ECU.

In the treatments enabling tax evasion, the supply curve is predicted to shift up by the amount of the effective tax, which depends on the reporting behavior of the traders and can differ from the 40 ECU nominal tax. In Seller only, the seller avoids the 40 ECU tax for the non-reported units but faces a 10% risk of being audited and having to pay the 40 ECU tax and a 40 ECU fine for the non-reported units. Since the expected monetary benefit of not reporting a unit, 40 ECU, exceeds the expected monetary cost, 8 ECU, money-maximizing and risk-neutral sellers do not report any of their sales. This implies that the expected effective tax per unit is 8 ECU, and the market supply curve thus shifts up by 8 ECU (solid blue line in Fig. 1). Therefore, market clears with a quantity of 16 units and a price between 161 ECU and 167 ECU. If some sellers are not money-maximizing and risk-neutral, they may report (some of) their trades, which implies that the expected effective tax is greater than 8 ECU. Our pre-registered hypothesis for Seller only treatment thus merely states that the market outcomes are between those in No tax and Automatic treatment.

[14] The complete post-experimental questionnaire is included in the pre-analysis plan which was submitted at the Open Science Framework before conducting any sessions (Kotakorpi et al. (2018)). The pre-registration also contains a design that was modified after pre-registration: A treatment in which the audit probability would be based on the effective audit probability observed in another treatment. The treatment design was modified after running the first treatments because the observed effective audit rate was so high so as to make the treatment redundant. The original treatment design was adapted and implemented as the Seller + BuyerC treatment. A treatment called "Social norm identification" was omitted completely and the data was never gathered due to restrictions caused by the Covid19 pandemic.

[15] We pre-registered predictions on market prices and quantities in treatments No tax, Automatic, Seller only, and Seller + Buyer. Predictions for treatment Seller + BuyerC were not pre-registered because the treatment was contrived only after the pre-registration form was submitted. Our original plan of running a treatment, where we exogenously impose audit rates implied by buyer-reporting, would have been meaningless as seller reports were already truthful. We thus decided to study the Seller + BuyerC -treatment where we expected higher buyer-reporting cost to yield the interim implied audit probability that we were expecting to have in the Seller + Buyer initially. This would allow us to study sellers' reactions to such imperfect third-party information, as was the original plan.

Table 2
General descriptive statistics, periods 11–25.

	No tax		Seller only		Seller + BuyerC		Seller + Buyer		Automatic
Reporting rate									
Sellers	-		28%	< ***	64%	< ***	88%		-
Buyers	-		-		43%	< ***	81%		-
Price									
Theoretical prediction	[158,162]		[161,167]		[161,167]		[161,182]		[178,182]
Observed	159.0	< ***	169.1	<	172.6	< *	177.7	<	177.7
Units sold									
Theoretical prediction	17		16		16		[13,16]		13
Observed	17.4	> ***	15.1	>	14.0	> ***	13.1	>	12.8
Observed effective tax									
Theoretical prediction	-		8		8		[8,40]		40
Observed	-		15.4		36.1		41.0		40.0
Tax incidence									
Theoretical prediction	-		50%		50%		50%		50%
Observed	-		62%	≠	37%	≠	46%	≠	47%

Notes: The table summarizes predicted and observed market outcomes, and observed reporting behavior. The reported figures are averages of market level means across periods 11–25 (see Table A.5 for market level means). Theoretical prediction refers to the predicted outcomes derived formally in Appendix C. Reporting rate of sellers (buyers) is the total number of trades reported by the sellers (buyers) divided by the total number of trades. Mean price is the mean price over all trades. Mean units sold is the mean number of units sold per period. Incidence of the observed effective tax is calculated by dividing the difference between the mean price in a given market and mean price in No tax by the observed effective per-unit tax in the corresponding market. Effective per unit tax refers to the total of all taxes and fines paid divided by the number of trades. The greater than and smaller than signs specify the direction of a one-sided Wilcoxon rank sum test and unequal sign indicates a two-sided test. Tests are conducted using market level means across relevant periods. The stars designate conventional levels of statistical significance (exact p-values, * $p<0.1$, ** $p<0.05$, *** $p<0.01$).

In the Seller + Buyer treatment, reporting is free to buyers, making them indifferent between reporting and not reporting. We consider the two boundary cases, assuming first that all buyers report all their trades. This implies that if a seller under-reports there is an audit with probability 80%, and hence the expected effective per-unit tax is 64 ECU ($0.8 \cdot 80$ ECU). Money-maximizing and risk-neutral sellers thus report all their sales, implying a 40 ECU upward shift in the market supply curve, just as in the Automatic treatment. Therefore, market clears with the quantity of 13 units and price between 178 ECU and 182 ECU. By contrast, assuming that all buyers report zero trades, evading sellers are subject only to the baseline 10% audit probability, making evasion worthwhile for money-maximizing, risk-neutral sellers. Markets are thus predicted to clear with a quantity of 16 units and a price between 161 ECU and 167 ECU, exactly as in the Seller only treatment. In practice, if some buyers report truthfully and others do not report any purchases, the expected effective tax is between 8 ECU and 40 ECU, and the market outcomes thus lie between those in the Automatic and Seller only treatments. This is also the prediction we pre-registered for the Seller + Buyer treatment.

In the Seller + BuyerC treatment, reporting is costly to buyers, which implies that money-maximizing buyers report zero trades. Facing no risk for an increased audit rate, money-maximizing and risk-neutral sellers evade fully, just as in the Seller only treatment. Therefore, market clearing price and quantity correspond to those in the Seller only treatment. However, if some buyers report truthfully despite the cost, the expected audit probability is above the 10% baseline rate and, accordingly, the expected effective tax per unit above 8 ECU. This implies that the market clearing price and quantity would be between those in the Seller only and Seller + Buyer treatments. Our pre-registered predictions are summarized in Table 1 (see Kotakorpi et al. (2018)).

5. Results

Table 2 displays a summary of predicted and observed market outcomes and reporting behavior in periods 11–25, when subjects have had the opportunity to learn and adapt their behavior. We discuss these results in the corresponding subsections below.[16]

5.1. Reporting rate

The left panel of Fig. 2 displays the evolution of the sellers' compliance rates by treatment (see Figs. B.10, B.11 and B.12 for evolution of reporting rates per individual market). It is evident that tax compliance crucially depends on the reporting mechanism (Kruskal-Wallis test p < 0.01 for all treatments jointly and for all pairwise comparisons, see also Fig. 3 and Table. A.14). In Seller

[16] We base our formal analysis on non-parametric tests based on ranks using market level means over periods 11–25 thus obtaining 6 independent observations per treatment. The market level means over periods 11–25 (1–25) are reported in Table A.5 (A.6). We report robustness checks of our main results using periods 1–25, 9–25, 10–25, 12–25 and 13–25 in Tables A.7, A.8 and A.9 Appendix A.

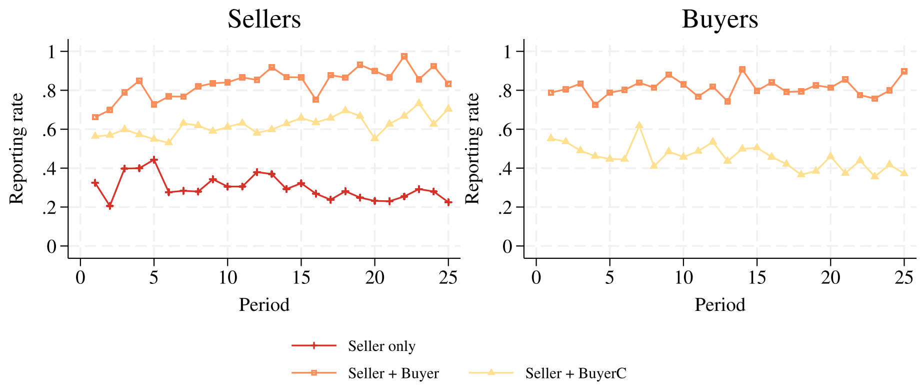

Fig. 2. Evolution of seller and buyer reporting rates by treatment. The figure plots the evolution of sellers' and buyers' reporting rates by treatment in periods 1–25.

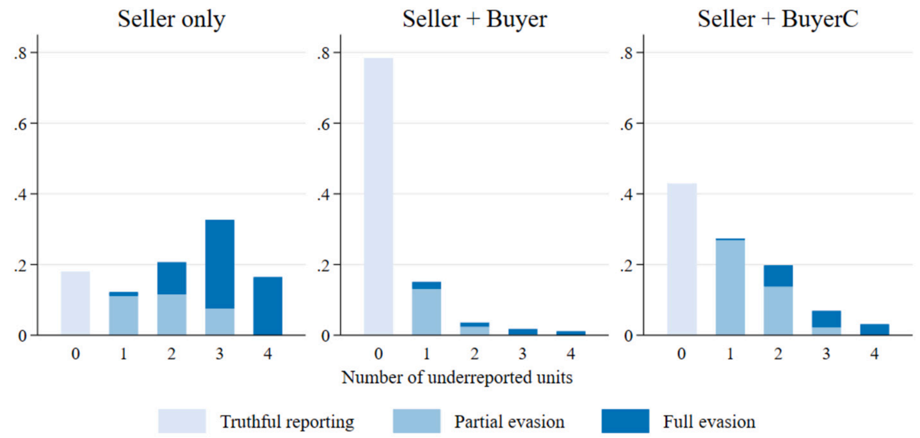

Fig. 3. Under-reporting. The Figure displays sellers' under-reporting by 0,1,2,3,4 units in SELLER ONLY, SELLER + BUYER and SELLER + BUYERC over periods 11-25.

ONLY sellers' compliance rate is 28% – significantly above the zero rate predicted by risk-neutral self-interest.[17] Compliance rates are even higher when there is buyer reporting: when buyer reporting is costless (in SELLER + BUYER), sellers report on average 88% of their trades in periods 11–25. When reporting is costly to the buyers (in SELLER + BUYERC), sellers' compliance rate is 64% on average in periods 11–25, which is well above the self-interested risk-neutral equilibrium prediction of 0%.

An explanation for the high compliance rates in SELLER + BUYER and SELLER + BUYERC can be seen in the right panel of Fig. 2. In SELLER + BUYER, buyers seem to break their indifference between reporting and not reporting in favor of truth-telling. Buyer reporting rate in periods 11–25 is nearly truthful, about 81% of the bought units are reported. 77% of the buyers' reports in that treatment are truthful and 15% are zero-reports. Buyers' reporting rate in SELLER + BUYERC is, as expected, significantly lower than in SELLER + BUYER, but also much higher than the zero reporting rate predicted by self-interest: 43% of the bought units are reported in periods 11–25, and 40% of the buyer reports are truthful whereas 50% are zero-reports.

The high buyer reporting rates in SELLER + BUYER and SELLER + BUYERC result in an environment where sellers have material incentives to generally report truthfully as well. In Appendix C where we formally derive the predictions, we also derive the expected audit probabilities conditional on buyer reporting rates (see Equation (3)). Substituting the observed average buyer reporting rates $p_{SB} = 0.81$ and $p_{SBC} = 0.43$ into Equation (3) yields audit probability schedules as a function of the number of sold units (row) and reported units (column) shown in Table 3.

The audit probabilities in SELLER + BUYER reported in Table 3 imply that when the seller is self-interested and risk-neutral and has correct expectations of buyer reporting, the only cases where under-reporting is optimal is when three or four units have been produced ($s_{i,t} = 3$ or $s_{i,t} = 4$). In that case, unit under-reporting is optimal, i.e. $r^*_{i,t} = 2$ and $r^*_{i,t} = 3$, respectively. When one or two units are supplied, truthful reporting is optimal.[18]

The audit probabilities for SELLER + BUYERC in Table 3 are such that, for a self-interested and risk-neutral seller with rational expectations, it is optimal to under-report by one unit if the amount produced is three or less, and by two units if it is four. Further

[17] For comparison, Doerrenberg and Duncan (2019) observe a compliance rate of about 7% in a closely related treatment.

[18] The predicted quantities supplied and optimal reporting behavior for a selfish risk-neutral seller by treatment and cost profile are shown in Table A.2 in Appendix A (see also Table A.3 for predictions for a fully compliant seller).

Table 3
Effective audit probability in Seller + Buyer and Seller + BuyerC.

Seller + Buyer

		Reported units ($r_{i,t}$)				
		0	1	2	3	4
Sold units ($s_{i,t}$)	0	10%				
	1	67%	10%			
	2	77%	56%	10%		
	3	80%	73%	47%	10%	
	4	80%	78%	68%	40%	10%

Seller + BuyerC

		Reported units ($r_{i,t}$)				
		0	1	2	3	4
Sold units ($s_{i,t}$)	0	10%				
	1	40%	10%			
	2	57%	23%	10%		
	3	67%	38%	16%	10%	
	4	73%	50%	25%	12%	10%

Notes: Effective audit probability in Seller + Buyer and Seller + BuyerC as a function of the number of sold units ($s_{i,t}$) and reported units ($r_{i,t}$) given the observed average buyer reporting rate p_{SB} = 0.81 in Seller + Buyer and p_{SBC} = 0.43 in Seller + BuyerC in periods 11–25.

under-reporting is deterred by the high implied marginal effect of expected taxes and fines which is due to the negative effects on infra-marginal units.

Notice that this yields a considerably sharper prediction for seller behavior in Seller + Buyer than the fairly general bounds identified in section C.1. In Seller + BuyerC, the thus received prediction deviates from the earlier prediction, as buyers do not entirely refrain from reporting their purchases as predicted by self-interest.

It turns out that the observed seller reporting behavior in Seller + Buyer and Seller + BuyerC mostly coincides with optimal responses to the audit rates (see Figs. 2 and 3), apart from reports being overly truthful. In Seller+Buyer, about 78% of sellers' reports are truthful when 40% are predicted to be truthful (see Table A.2 in Appendix A for predicted number of trades and reported units for a standard risk-neutral seller). Only 6% of reports under-report more than one unit, which is very close to the predicted 0%. Yet, this sub-optimal under-reporting is sufficient to tilt the expected effective tax above the 40 ECU which would result with truthful reporting. Moreover, under-reporting is more common in Seller + BuyerC than in Seller + Buyer, as predicted. Furthermore as predicted, under-reporting by one unit is the most common form of under-reporting in Seller + BuyerC unlike in Seller only (see Fig. 3). Considering individual seller cost profiles (Table A.1), the optimal responses to buyer-reporting predict that no seller should be truthful, but under-report by one unit, and no-one should under-report fully (see Table A.2). By and large, these predictions are borne out in the data. Yet, some sellers report too many trades (see Fig. B.15 in Appendix B). It therefore appears that the mere possibility of buyer reporting, even if third-party information is not effectively supplied by buyers, may have a disciplining effect on seller reporting behavior.

The main pattern not captured by these predictions, however, is that even though reporting responds to deterrence in a standard way, the level of reporting tends to be overly truthful across treatments and player roles: 40% of the buyer reports and 42% of the seller reports are truthful in Seller+BuyerC, 28% of the seller reports are truthful in Seller only, all inconsistent with the self-interested and risk-neutral equilibrium prediction; moreover, reporting is more consistent with the truthful than the payoff-dominating equilibrium (whereby indifferent buyers report truthfully) in Seller+Buyer.

We will return to this issue in Section 6 where we compare explanations that account for these reporting patterns and our other results.

5.2. Market prices and quantities

The evolution of mean market prices and quantities is displayed in Fig. 4 (see Figs. B.5–B.9 for evolution of prices and quantities per individual market). The dotted lines indicate the predicted market price and quantity in the No tax and Automatic benchmark conditions. As can be observed, our results are largely consistent with these predictions, especially in later periods when market participants have had ample opportunity to learn and adapt (see Table 2).[19]

[19] In particular, the mean price in No tax in periods 11–25 is not statistically significantly different from 160 ECU, the mid-point of the equilibrium price range. The mean quantity sold, 17.4 units on average, is slightly higher than the predicted 17 units. Similarly, the mean price in Automatic is not statistically significantly different from 180 ECU in periods 11–25. Furthermore, the mean quantity per period is not statistically significantly different from the predicted 13. (Wilcoxon signed rank test, $p > 0.1$, except for quantity in No tax treatment.) The exact p-values and treatment level means of prices, median prices and quantities for different

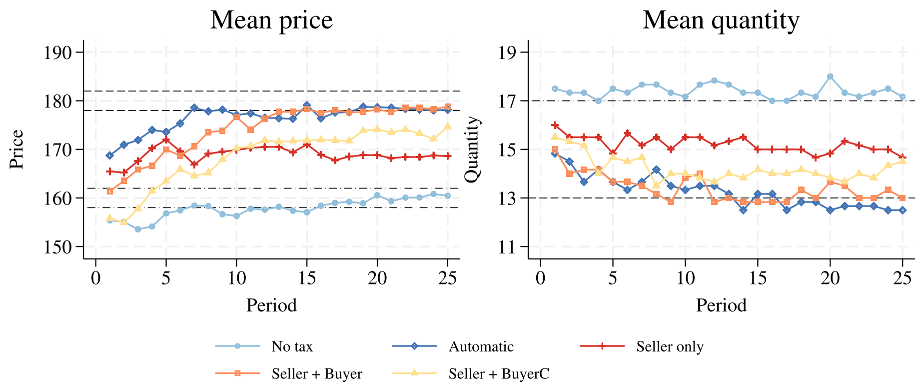

Fig. 4. Evolution of mean prices and quantities. The figure plots the evolution of mean market prices and quantities sold by treatment in periods 1-25. The mean price is the average of the mean price in a given market period, and mean quantity is the average of the number of units traded in a given market-period. The dotted lines denote the predicted market price intervals and quantities in No tax ([158,162] with 17 units traded) and Automatic ([178,182] with 13 units traded) benchmark conditions.

In line with our pre-registered predictions, market prices are significantly lower and quantities traded significantly higher in Seller only than in Automatic.[20] However, prices are higher and quantities lower than implied by the 8 ECU upward shift of the supply curve predicted by self-interested risk neutrality, suggesting a potential role for behavioral factors (see Appendix C).

With third-party information about transactions in Seller + Buyer, market prices and quantities are very close to those in Automatic. This result is consistent with the observed high reporting rates and thereby the compliant equilibrium prediction (see Section C.2).

Also in Seller + BuyerC where reporting is costly to the buyers, the findings are in line with the predictions. First, the mean price is lower than in Seller + Buyer (and Automatic) in periods 11–25. Furthermore, the mean quantity traded per period is higher than the mean quantity in Automatic and Seller + Buyer. As discussed in Section 5.1, costly reporting decreases the amount of available third-party information. As a consequence, sellers evade more, pushing down market prices. However, lower prices only translate into a weakly higher market activity in terms of quantities traded. Comparing Seller + BuyerC to Seller only, we find that there is no significant difference in the mean price nor mean quantities traded in periods 11–25, in line with the non-compliant equilibrium prediction. However, given that both seller and buyer reporting rates are higher than predicted, the non-significant difference in market activity is noteworthy.

5.3. *Tax incidence*

Fig. 5 illustrates the share of the observed effective unit tax borne by buyers in each market in the four treatments with taxes. The market level means and treatment averages are reported in Tables A.5 (periods 11–25) and A.6 (periods 1–25).[21]

Considering the main periods 11–25 buyers bear 46.8% (46.1% in periods 1–25) of the tax burden in Automatic, 62.4% (65.9%) in Seller only, 45.7% (38.4%) in the Seller + Buyer and 37.5% (28.0%) in Seller + BuyerC. The incidence therefore is strikingly close to the theoretical 50-50 prediction in Automatic and Seller + Buyer treatments. Moreover, there is very little variation between markets in those treatments. In contrast, the incidence varies substantially in Seller + BuyerC (from 17.6% to 56.1%) and even more in Seller only (from -0.04% to 105.6%) between markets.In Seller + BuyerC sellers appear to struggle to shift the tax burden onto buyers, whereas in Seller only they are able to shift a notably large share of the effective tax onto buyers in four out of six markets.[22]

subsets of periods (periods 11–25, 1–25, 9–25, 10–25, 12–25 and 13–25) are reported in Tables A.7, A.8 and A.9. The results are robust with respect to the choice of periods; an exception is the mean price in Automatic across periods 1–25 which differs from the predicted price (exact p-value = 0.0464), but does not differ when considering the subsets of later periods in which markets have converged.

[20] We thus qualitatively confirm the corresponding result in Doerrenberg and Duncan (2019).

[21] We calculate the effective unit tax as the total of the taxes and fines paid in a given market divided by the total number of units sold in a given market. The incidence of the effective unit tax is calculated by subtracting the mean price in the No tax treatment from the mean market price and dividing by the effective tax. We calculate the incidences using data from periods 11–25 and periods 1–25. The evolution of the *nominal* per unit tax by treatment is shown in Fig. B.14. Treatment level descriptive statistics on taxes and fines paid are reported in Tables A.16 and A.15.

[22] The incidence is not statistically significantly different from the predicted 50% in any of the treatments (exact p-value > 0.1 for two-sided Wilcoxon signed rank test using 6 market level means as observations). We test for differences between treatments using Wilcoxon rank sum tests of differences in means and report the approximate and exact p-values, which take into account the small number of observations, in Table A.10. The WRS test assumes equality of variances in the distributions, and therefore we interpret the results with caution. In Seller only there is one anomalous market (session number 15) in which the mean price is strikingly low, resulting in a lower expected effective incidence than in any other market in our experiment. Conducting the WSR test without the outlier market suggests that the incidence in Seller only indeed is statistically different from Seller + BuyerC (exact p-value =0.030), and from Seller + Buyer and Automatic (exact p-value = 0.082 for both). Dropping observations (markets) one at a time also changes the test result between Seller + Buyer and Seller + BuyerC (when dropping market 28, the WRS-test results in a weakly statistically significant difference, exact p-value = 0.052), but otherwise does not change the results. Testing for the difference using incidence aggregated across all periods (1–25) indicates that incidence in Seller only differs from the other treatments (exact p-value <0.1 in

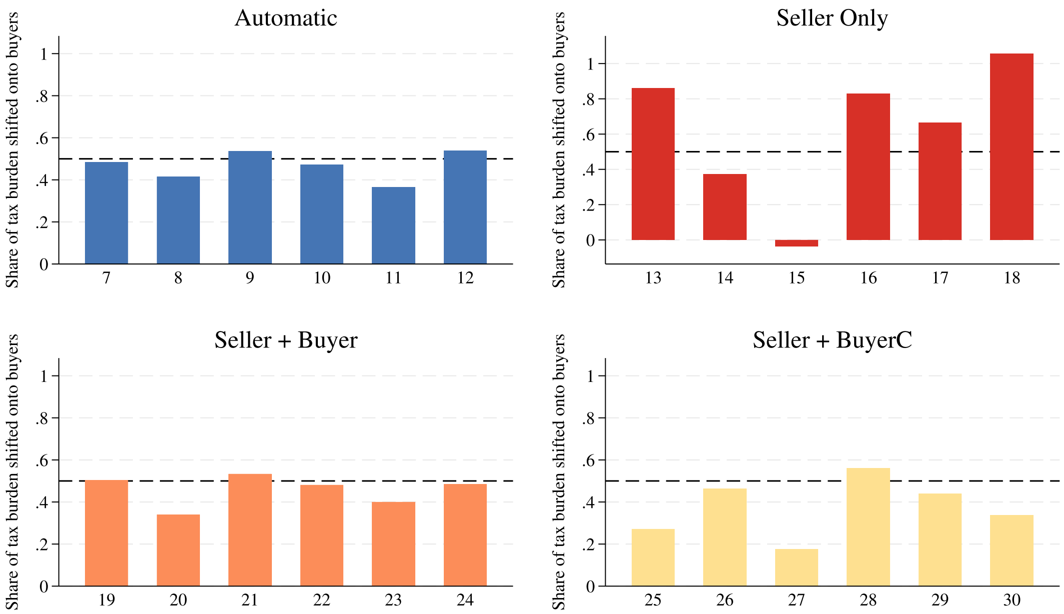

Fig. 5. Incidence of observed effective tax, periods 11-25. The figure displays the incidence of the observed effective unit tax by market (numbered from 7 to 30) in treatments AUTOMATIC, SELLER ONLY, SELLER + BUYER and SELLER + BUYERC. Observed effective unit tax is defined as the total of taxes and fines paid divided by the number of trades in a given market. Share of the tax borne by buyers is calculated as the difference in mean price in a market minus the mean price in the NO TAX treatment divided by the observed effective unit tax. The figure shows incidences using data from periods 11-25 (see Fig. B.13 for figures using data from all periods).

Fig. 6 shows the frequencies of the average trading prices of sellers by their evasion decision in SELLER ONLY (panel A), SELLER + BUYER (panel B) and SELLER + BUYERC (panel C). The histograms report frequencies within each treatment. In SELLER ONLY, a notable proportion of fully evading sellers sell at the NO TAX predicted equilibrium price, but an even larger share of fully evading sellers sell at higher prices. In other words, many sellers in SELLER ONLY evade fully, yet set their prices as if they were to report truthfully. The price setting and reporting behavior of this group of sellers largely explains why sellers are able to shift a significant share of the tax burden onto buyers.

6. Alternative mechanisms

We next discuss different mechanisms that may explain our non-standard incidence result. We first argue that our findings are not explained by risk aversion. We then turn to behavioral models of lying behavior. In particular, we apply a model that builds on Gneezy et al. (2018) and Abeler et al. (2019), and incorporates both lying costs and image concerns into decision-making. We also consider a simpler nested model with lying costs only. Finally, we discuss whether our findings are consistent with a model of intrinsic generosity or philanthropy towards the state.

Doerrenberg and Duncan (2019) argue that risk aversion may explain why evasion leads to higher prices than predicted by the standard model: sellers may seek compensation for the risk related to evasion and shift more of the effective tax burden on buyers in the treatments with tax evasion. However, it is not clear a priori that risk aversion should lead to higher posted prices per unit. First, one key source of risk is that of failing to sell. This points to risk-averse sellers charging lower, not higher prices, to increase the probability of trade.[23] In addition, in audit risk treatments, greater risk aversion is predicted to lead to a higher reporting rate. In our data, risk attitudes are not predictive of the average trading prices, number of trades, or the reporting rate. Risk aversion is negatively correlated with trading prices and positively correlated with the reporting rate, while neither of these correlations is statistically significant (see Tables A.12 and A.13 in Appendix A). Further, the theoretical prediction of the effect of risk aversion on average prices is considerably more complicated in treatments with endogenous audit risk (SELLER + BUYER and SELLER + BUYERC) and it is certainly not clear that the effect on price is positive.[24]

all comparisons) and SELLER + BUYER and SELLER + BUYERC differ (exact p-value = 0.015) (see Table A.11). We also report test results from Kolmogorov-Smirnov tests testing the differences in the distributions. The results are similar, but the small number of observations limits the scope of interpretation.

[23] In the NO TAX and AUTOMATIC treatments where audit risk is not an issue, this is the only source of risk, and we obtain an unambiguous prediction that risk-averse buyers should charge lower prices.

[24] In those treatments risk aversion increases the reporting rate. Depending on the buyer reporting rate this may, however, have a positive or negative effect on expected marginal cost (i.e. expected effective tax). Moreover, it is not clear that a risk-averse seller is in a position of charging higher prices when competing with less risk-averse or risk-neutral ones.

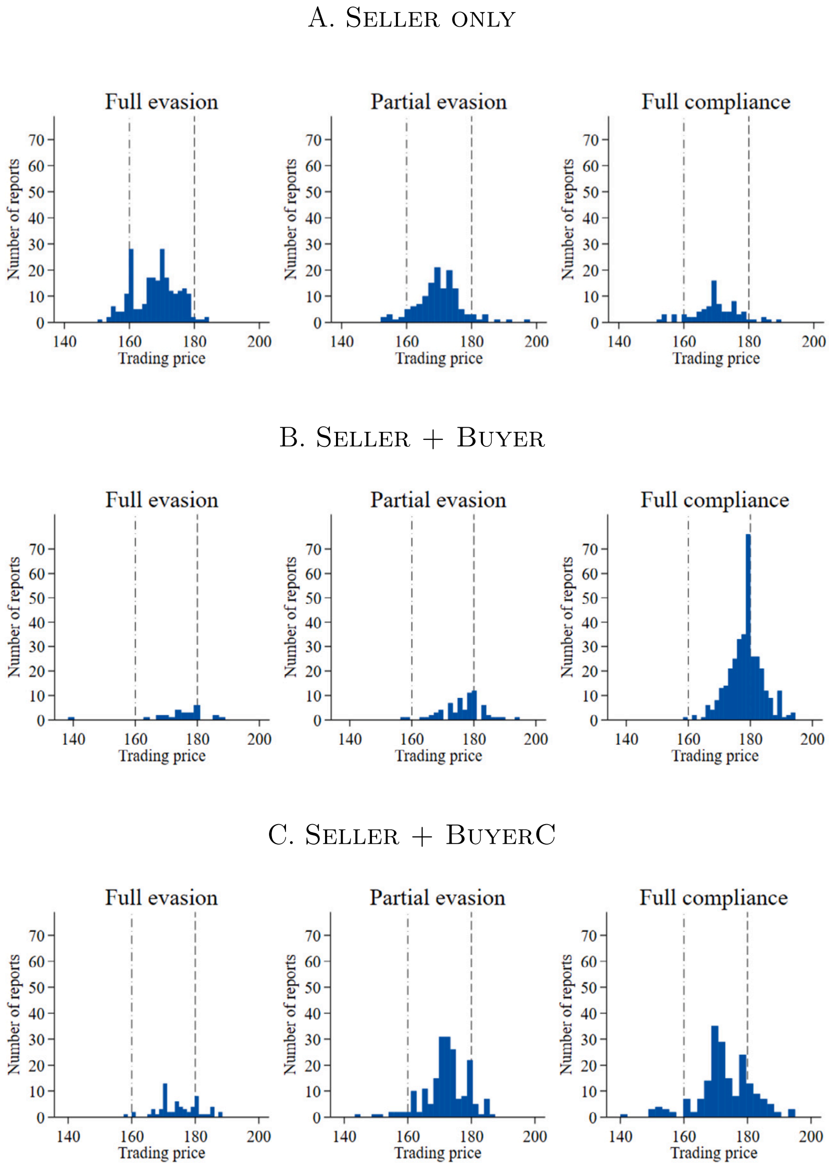

Fig. 6. Trading prices by seller's reporting decision. Frequencies of average trading prices by sellers' reporting decisions (full evasion, partial evasion or full compliance) within a treatment using seller-period level data from periods 11–25. Vertical lines indicate the (midpoint of) predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Second, we turn to behavioral models of lying. In Appendix D, we apply the lying aversion model of Gneezy et al. (2018) to our setup. We show that our findings across the board are consistent with the model: the model can explain the non-standard effective tax incidence patterns, as well as our findings on reporting and market behavior. In the model, sellers take advantage of evasion opportunities by (privately) under-reporting sales, but due to image concerns do not want to reveal themselves as evaders by offering low prices in the market stage.[25] Hence, even though average prices are somewhat lower under evasion, they do not fully reflect the lower monetary costs of the evaders. This can explain why incidence is heavily distorted towards buyers in the Seller only treatment where evasion is prevalent.

A third potential explanation for the patterns we observe is session-level variation in the distribution of moral motivation among the active sellers. Such an explanation hinges on substantial heterogeneity and variation in lying costs between sessions, but does

[25] In an open-ended question in the post-experimental questionnaire, about a fifth of buyers in the two treatments with buyer-reporting mentioned (without any solicitation along these lines) that they expected an association between tax evasion and (low) prices. This lends indirect support to our conjecture that image concerns may have played a role in the market stage.

not require image concerns to play a role. From the perspective of an individual tax-evading seller in the Seller only treatment, the degree of competition in the marketplace is determined by how many morally motivated sellers there are who are willing to report truthfully despite the material incentives not to do so. To be more precise, let us consider two extreme alternatives. First, if all sellers report truthfully, then they are unable to compete fiercely and the prediction in such a session coincides with that of Automatic. Second, if all sellers evade and report none of their produced units, then competition is fierce (in the absence of image concerns), and the prediction is closer to No Tax. This kind of heterogeneity in moral motivation can generate the variation in prices we observe in the Seller Only treatment. To yield a prediction consistent with our data, there should not only be sessions of the extreme type as described above, but also sessions with a specific mixture of sellers who charge high prices and report truthfully (lying-averse types) and sellers who charge similar prices but do not report their sold units (standard types). In particular, there should be sessions with sufficiently many sellers with lying aversion so that prices charged by the tax-evading sellers will not be competed down.[26] The observed considerable fraction of tax-evading sellers who charge price equally high as truthful types should come from such sessions.

This explanation is consistent with the evidence from other treatments, too, because in Automatic, Seller + Buyer, and Seller + BuyerC, the audit probability of an evading seller is so high that even self-interested sellers should find it optimal to report most or all of their sold units.

A fourth candidate to explain the data is a model of intrinsic generosity or philanthropy towards the state.[27] However, we argue below that this type of a model is not consistent with the patterns in our data. As in the case of the model combining lying aversion and willingness to appear honest, the intrinsic motivation of generosity could be combined with a motive to appear generous. Such a model could explain the higher-than-predicted reporting rates by the sellers in Seller only. Sellers who are motivated by generosity towards the state report their trades truthfully. In the market stage, these sellers should charge lower prices, however, and thereby personally contribute to the state finances, rather than passing the expenses of their generosity on others (i.e. buyers).[28] Moreover, maximizing state revenue would require maximizing the probability of trade to expand the tax base. Increasing the probability of trade implies charging lower, not higher prices. Thus, unlike lying aversion, the generosity motive pushes prices down.

Now, sellers wanting to *appear* generous (without necessarily being generous) at the market stage must also charge low rather than high prices in order to pool with the generous sellers in equilibrium. Thus, in contrast to models with lying aversion, the generosity model appears to predict low prices. Furthermore, a model with generosity cannot directly explain the positive reporting rates by the buyers in the Seller+BuyerC treatment.[29] A buyer paying the reporting fee does not contribute to the state finances, and thus the truthful buyer-reporting, when costly, requires another explanation, such as lying aversion.

Based on the above discussion we can conclude that, whereas generosity or risk aversion models cannot explain our findings, lying aversion models are consistent with our data. As our design is not built to explicitly test the different versions of the lying aversion model, we cannot exactly pin down which of them provides the most accurate description of individual motivations in our context. Abeler et al. (2019) conclude that a model with both lying costs and image concerns is required to explain behavior in the earlier experimental literature, and we explicitly show how such a model can be modified to analyze tax evasion, with separate tax reporting and market phases. Future work may build on this contribution to test the model explicitly. In our particular case, also a simpler model with lying costs only appears sufficient to explain our experimental findings. A key difference is that in the model with both lying costs and image concerns, both of these forces work towards keeping prices higher than in the standard model; while in a model with competition and heterogeneous lying costs, only the latter mechanism is operational.

7. Discussion

Some of the previous literature interprets the effectiveness of third-party information in deterring evasion as evidence that moral motivation is not an important factor in tax evasion behavior: according to this interpretation, the key factor appears not to be that individuals are unwilling to cheat, but that they are unable to cheat due to the prevalence of third-party information (Kleven et al., 2011).

We find, in line with that literature, that the way in which reporting behavior responds to deterrence is largely in line with the standard, self-interested theory. However, our evidence suggests that intrinsic costs related to lying and/or image concerns may nevertheless be important to many. Such costs imply that the level of tax evasion is lower on average than predicted by a standard, self-interested model. Intrinsic costs may also have important repercussions on market behavior and outcomes, leading to departures from the incidence results of standard theory.

Our findings may at first sight appear puzzling in this regard: the comparative statics of reporting behavior appear consistent with the standard self-interested model, while market behavior does not. We show that the apparently conflicting interpretations can be reconciled within a single model framework with moral motivations in the form of lying costs and image concerns. Reporting

[26] Standard types could also sell more units (due to avoiding the tax). This way their marginal cost would coincide with that of the lying-averse types (who have lower production cost but who pay the tax). This would further contribute to prices not being competed down. This would generate the higher of the two modes in top-left panel of Fig. 6 coinciding with the price mode of the fully compliant (top-right panel) and partially compliant sellers (top-middle panel).

[27] This possibility was raised by an anonymous referee and was prompted by the fact that the experimental tax revenue was donated to the state even if there is no legal obligation to do so.

[28] Especially given that evidence in the literature favors a warm-glow form of generosity, or impure altruism, over pure altruism (Andreoni, 2006; Tonin and Vlassopoulos, 2014; Imas, 2014).

[29] Again, the warm-glow version of the generosity model is particularly explicit about this pattern.

behavior responds to deterrence as in the standard model, albeit levels of compliance are higher than in the standard model due to the presence of lying costs. Image concerns on the other hand have no effect on *reporting* behavior, as reporting decisions are not observable to other market participants. However, image concerns do play a role when it comes to behavior in the marketplace. The desire to appear honest implies that evasion in the reporting stage is not fully reflected in lower prices in the trading stage, thus giving rise to the non-standard incidence result in the presence of evasion. Price competition with heterogeneous lying costs may lead to a similar phenomenon.

Another factor that affects the implications of third-party reporting is whether the relevant third parties have incentives to provide such information in the first place. Existing discussions of third-party reporting typically take the existence (or the lack) of the required information as given. In our experiment, if buyers take into account the potential indirect effect of tax evasion on prices, they should not provide reports that enable stricter tax enforcement. Nevertheless, we find that introducing buyer reporting has a very strong disciplining effect on evasion. Buyers and sellers are thus unable to tacitly collude on an outcome with a lower level of reporting, even though it would be in their joint monetary interest. We show in our theoretical model, that also this finding can be explained by moral motivation: even very little (some) intrinsic motivation to report honestly is sufficient to destabilize the collusive equilibrium when buyer reporting is costless (costly). Making reporting more costly potentially facilitates such collusion, and indeed, buyers are less likely to report when reporting comes with a cost. Compliance of sellers decreases as well, but not by as much.

From a policy perspective, an interesting conjecture from these results is that taxpayer awareness of the mere *possibility* of the existence of third-party information, even if it is not effectively supplied by buyers, may be a fairly effective deterrent on tax evasion.

Tax administrators have traditionally been quite secretive about the type of information used in tax enforcement. Our findings point to a tentative policy implication, namely that providing information on the *types* of third-party information available to the tax authority could be an effective way of deterring tax evasion. (Notifying taxpayers of the exact information held by the tax authority on the other hand may in some cases backfire, a result found by Slemrod et al. (2017) and Carrillo et al. (2017) in a natural experiment setting.)[30] Naturally, providing information should be made as cheap as possible to the third party. Costly information provision has two downsides: the reporting cost constitutes an efficiency loss, and may facilitate collusion between the buyer and seller to jointly evade taxes.

Our results raise some interesting directions for future research. First, it is of importance to further carefully test the behavioral theory which is found widely consistent with the observations in the present paper, as well as earlier experimental evidence on lying behavior (Abeler et al., 2019), in the tax evasion context. Second, a potentially interesting extension to our analysis includes studying a non-anonymous setting. In this case, the effects of moral motivation – in particular image concerns – are likely to be more pronounced. Third, allowing for communication may have nuanced effects as it facilitates collusion but may on the other hand reinforce the impact of image motivation.

Finally, in our study, market agents are randomly assigned to market institutions and seller or buyer roles. To increase the external validity of the exercise and to understand the role of selection on competition and tax evasion for market outcomes, it would be of interest to endogenise the entry to markets. As the evidence in Halliday et al. (2021) suggests, it is conceivable that agents without moral costs are likely to select into markets where they benefit from their comparative advantage, whereas agents prone to lying aversion, self-select into markets where deterrence is effective or lying is not necessary.[31] The implications of selection are not straightforward, as presence of certain behavioral types will generate externalities on others, and each type is likely to select into and away from certain kind of populations of types and institutional environments: intriguing open questions include, for example, whether groups of dishonest sellers and buyers self-select into the same markets, enabling collusion on a high-evasion, low-price outcome; or whether dishonest and honest sellers coexist in the market, thus potentially conferring a competitive advantage to dishonest agents.

Declaration of competing interest

The authors declare that they have no relevant material or financial interests that relate to the research described in this paper.

Data availability

Data will be made available on request.

Appendix A. Additional tables

[30] Okat (2016) shows in a theory model that it may be optimal not to use all available information in tax enforcement, in the sense that randomness in audit rules may be optimal to prevent learning by evaders; see Alm et al. (1992) for a related argument plus caveats. A few previous lab experiments have discussed the impact of ambiguity – not having precise knowledge of objective audit probabilities – on tax evasion, and according to a recent review of tax experiments those papers found mixed results (Malézieux, 2018). Agranov and Buyalskaya (2022) show that shrouding information concerning a fine distribution may be an effective way to deter lying.

[31] Harju et al. (2023) find with field data on firm tax audits in Finland, that audits lead to an increase in the likelihood of firm exits, in particular for firms that were found to be non-compliant.

Table A.1
Demand and supply schedules.

Buyer	Value 1	Value 2	Value 3	Value 4
1	232	207	182	177
2	212	202	192	152
3	227	222	167	157
4	242	197	172	147
5	237	217	187	162
Seller	Cost 1	Cost 2	Cost 3	Cost 4
1	88	113	138	143
2	108	118	128	168
3	93	98	153	163
4	78	123	148	173
5	83	103	133	158

Notes: Costs and values were randomized into sets of four. The sets of four costs/values were randomly assigned to traders at the beginning of each experimental session.

Table A.2
Predicted quantities by treatment and cost profile for a standard type seller.

$\theta = 0; \alpha = 0$	C1	C2	C3	C4	C5
NT	4	3	3	3	4
A	2	3	2	2	3
SO	4(0)	3(0)	3(0)	3(0)	4(0)
SB	3(2)	3(2)	2(2)	2(2)	3(2)
SBC	2(1)	3(2)	2(1)	2(1)	2(1)

Notes: The table displays the predicted quantities sold (reported) by treatment and cost profile for a standard type seller using the observed mean price in a given treatment in periods 11–25. The predictions for Seller + Buyer and Seller + BuyerC are founded on the expected audit probability based on the average observed buyer reporting rates in periods 11–25 in the respective treatments.

Table A.3
Predicted quantities by treatment and cost profile for a fully compliant seller.

$\theta >> 0$	C1	C2	C3	C4	C5
NT	4	3	3	3	4
A	2	3	2	2	3
SO	2(2)	3(3)	2(2)	2(2)	2(2)
SB	3(3)	3(3)	2(2)	2(2)	3(3)
SBC	2(2)	3(3)	2(2)	2(2)	2(2)

Notes: The table displays the predicted quantities sold (reported) by treatment and cost profile for a fully compliant seller using the observed mean price in a given treatment in periods 11–25.

Table A.4
Summary statistics of demographic variables.

	Gender	Age	Finnish	Tax morale	Risk attitude	Generalized trust
No tax						
Mean	0.73	29.17	1	0.28	4.73	6.07
St. Dev.	-	9.46	-	-	2.09	2.31
N. of Subjects	60	60	60	60	60	60
Automatic						
Mean	0.75	26.77	0.93	0.62	5.02	6.33
St. Dev.	-	7.01	-	-	1.99	2.10
N. of Subjects	60	60	60	60	60	60
Seller only						
Mean	0.75	28.02	1	0.62	5.37	5.85
St. Dev.	-	8.09	-	-	2.45	2.44
N. of Subjects	60	60	60	60	60	60
Seller + Buyer						
Mean	0.62	27.57	0.97	0.57	5.18	6.87
St. Dev.	-	5.58	-	-	2.24	2.17
N. of Subjects	60	60	60	60	60	60
Seller + BuyerC						
Mean	0.75	27.25	0.95	0.52	5.20	6.78
St. Dev.	-	7.94	-	-	2.11	1.90
N. of Subjects	60	60	60	60	60	60
P-value	0.40	0.45	0.12	0.001	0.46	0.26

Notes: Reported are the mean characteristics of the five treatment groups. Gender is an indicator that is equal to 1 if the subject is female, Finnish is an indicator that is equal to 1 if the subject's native language is Finnish. Tax morale is an indicator that is equal to 1 if the subject reported that cheating on taxes is never acceptable. Risk attitude is the subject's reported willingness to take risks (0 = "not at all willing" to 10 = "very willing"), and generalized trust is the subject's reported propensity to trust other people (0 = "one can never be too careful with other people" to 10 = "most people can be trusted"). P-values are for χ^2 test, apart from Age for which the p-value is for Kruskal-Wallis test. For each test, the null hypothesis is that there are no differences between the five treatment groups.

Table A.5
Market level results, periods 11–25.

	Session number	Mean price	Median price	Mean quantity	Seller reporting rate	Buyer reporting rate	Effective per unit tax	Incidence of tax
No tax	1	157.0	155	17.5				
	2	152.7	155	17.3				
	3	160.5	160	18.1				
	4	157.9	160	17.3				
	5	163.4	164	17.3				
	6	162.4	165	16.9				
	avg.	**159.0**	**159.8**	**17.4**				
Automatic	7	178.5	180	12.7			40.00	0.48
	8	175.5	175	13.7			40.00	0.41
	9	180.5	181	12.7			40.00	0.54
	10	177.9	179	12.7			40.00	0.47
	11	173.6	175	12.7			40.00	0.36
	12	180.5	180	12.6			40.00	0.54
	avg.	**177.7**	**178.3**	**12.8**			**40.0**	**0.47**
Seller only	13	173.2	173	14.5	0.35		16.41	0.86
	14	165.3	165	16.3	0.32		17.05	0.37
	15	158.6	160	16.4	0.16		9.59	-0.04
	16	170.3	170	13.0	0.26		13.95	0.83
	17	173.3	172	14.5	0.35		21.57	0.66
	18	174.0	170	15.9	0.24		14.12	1.06
	avg.	**169.1**	**168.3**	**15.1**	**0.28**		**15.4**	**0.62**
Seller + Buyer	19	179.7	178	13.7	0.96	0.99	41.17	0.50
	20	173.0	173	14.1	0.82	0.80	41.13	0.34
	21	179.9	180	12.5	0.90	0.88	39.15	0.53
	22	179.7	180	12.2	0.78	0.74	43.06	0.48
	23	175.6	177	12.3	0.88	0.82	41.73	0.40
	24	178.2	180	13.9	0.92	0.65	39.62	0.48
	avg.	**177.7**	**178.0**	**13.1**	**0.88**	**0.81**	**41.0**	**0.46**
Seller + BuyerC	25	169.5	170	13.8	0.78	0.50	38.65	0.27
	26	174.5	177	13.3	0.70	0.36	33.80	0.46
	27	164.7	161	14.9	0.69	0.29	32.32	0.18
	28	179.9	180	13.5	0.55	0.28	37.24	0.56
	29	175.0	175	13.6	0.69	0.63	36.27	0.44
	30	172.0	170	14.7	0.45	0.54	38.55	0.34
	avg.	**172.6**	**172.1**	**14.0**	**0.64**	**0.43**	**36.1**	**0.37**

Notes: Table reports market level means and treatment averages of mean prices, median prices, mean quantities, seller reporting rates, buyer reporting rates, effective per unit taxes and incidence across periods 11–25.

Table A.6
Market level results, periods 1–25.

	Session number	Mean price	Median price	Mean quantity	Seller reporting rate	Buyer reporting rate	Effective per unit tax	Incidence of tax
No tax	1	157.3	155	17.6				
	2	148.6	150	17.1				
	3	159.8	160	17.9				
	4	157.1	160	17.2				
	5	162.4	163	17.4				
	6	162.1	162	17.1				
	avg.	**157.9**	**158.3**	**17.4**				
Automatic	7	173.9	180	13.4			40.00	0.39
	8	177.4	175	13.8			40.00	0.49
	9	181.0	180	12.9			40.00	0.58
	10	177.6	179	13.2			40.00	0.49
	11	172.5	172	13.4			40.00	0.36
	12	176.7	180	12.8			40.00	0.46
	avg.	**176.5**	**177.7**	**13.3**			**40.0**	**0.46**
Seller only	13	170.2	171	14.8	0.36		16.69	0.73
	14	166.5	166	16.3	0.34		17.00	0.51
	15	160.5	160	16.2	0.19		10.40	0.24
	16	170.3	170	13.7	0.27		16.33	0.77
	17	170.7	171	14.6	0.36		21.20	0.60
	18	175.1	172	15.7	0.28		15.37	1.11
	avg.	**168.9**	**168.3**	**15.2**	**0.30**		**16.2**	**0.66**
Seller + Buyer	19	179.0	177	14.2	0.89	0.94	42.71	0.49
	20	168.4	171	14.4	0.75	0.82	40.33	0.26
	21	177.2	180	12.8	0.88	0.84	41.25	0.46
	22	175.5	189	12.6	0.74	0.76	45.21	0.37
	23	169.4	174	12.6	0.86	0.84	41.77	0.27
	24	176.0	179	13.8	0.89	0.67	40.58	0.45
	avg.	**174.2**	**176.7**	**13.4**	**0.84**	**0.81**	**42.0**	**0.38**
Seller + BuyerC	25	166.4	170	13.6	0.74	0.53	38.11	0.22
	26	168.4	170	13.7	0.68	0.37	37.20	0.28
	27	160.4	160	15.2	0.63	0.31	36.64	0.06
	28	175.5	179	14.0	0.50	0.33	38.28	0.45
	29	171.9	171	13.6	0.70	0.69	37.89	0.37
	30	169.3	170	15.1	0.48	0.51	37.46	0.28
	avg.	**168.6**	**170.0**	**14.2**	**0.62**	**0.46**	**37.6**	**0.28**

Notes: Table reports market level means and treatment averages of mean prices, median prices, mean quantities, seller reporting rates, buyer reporting rates, effective per unit taxes and incidence across periods 1–25.

Table A.7
Robustness of treatment effects on mean price.

	No tax Predicted [158,162]	<	Seller only	<	Seller + BuyerC	<	Seller + Buyer	<	Automatic Predicted [178,182]
11–25									
	159.0 [p>0.1]	p = 0.0076	169.1	p = 0.1970	172.6	p = 0.0660	177.7	p = 0.4686	177.7 [p>0.1]
1–25									
	157.9 [p>0.1]	p = 0.0043	168.9	p = 0.6504	168.6	p = 0.0465	174.2	p = 0.1970	176.5 [p = 0.0464]
9–25									
	158.7 [p>0.1]	p = 0.0076	169.2	p = 0.2944	172.2	p = 0.0325	177.4	p = 0.4091	178 [p>0.1]
10–25									
	158.8 [p>0.1]	p = 0.0076	169.2	p = 0.2944	172.4	p = 0.0325	177.6	p = 0.4091	177.7 (p>0.1)
12–25									
	159.1 [p>0.1]	p = 0.0076	169.0	p = 0.2424	172.7	p = 0.0206	177.9	p = 0.4091	177.8 (p>0.1)
13–25									
	159.2 [p>0.1]	p = 0.0076	168.9	p = 0.1970	172.8	p = 0.0660	178.1	p = 0.4686	177.9 (p>0.1)

Notes: This table presents mean market prices per treatment using different subsets of market periods, and exact p-values from one-sided Wilcoxon rank sum tests and two-sided Wilcoxon signed rank tests. The rank sum test tests the inequalities between mean prices in treatments. The Wilcoxon signed rank test tests the equality of observed prices and midpoint of predicted prices in treatments No tax and Automatic (in squared brackets below the observed price). Tests are conducted using market level means across relevant periods.

Table A.8
Robustness of treatment effects on mean quantity.

	No tax Predicted 17	>	Seller only	>	Seller + BuyerC	>	Seller + Buyer	>	Automatic Predicted 13
11–25									
	17.4 [$p = 0.0625$]	$p = 0.0011$	15.1	$p = 0.1126$	14.0	$p = 0.1548$	13.1	$p = 0.5271$	12.8 [p > 0.1]
1–25									
	17.4 [$p = 0.0313$]	$p = 0.0011$	15.2	$p = 0.0509$	14.2	$p = 0.1548$	13.4	$p = 0.5314$	13.3 [p > 0.1]
9–25									
	17.4 [$p = 0.0625$]	$p = 0.0011$	15.1	$p = 0.1115$	14.0	$p = 0.1861$	13.2	$p = 0.4686$	12.9 [p > 0.1]
10–25									
	17.4 [$p = 0.0625$]	$p = 0.0011$	15.1	$p = 0.1115$	14.0	$p = 0.1861$	13.2	$p = 0.4892$	12.9 [p > 0.1]
12–25									
	17.4 [$p = 0.0625$]	$p = 0.0011$	15.0	$p = 0.1093$	14.0	$p = 0.1115$	13.1	$p = 0.5844$	12.8 [p > 0.1]
13–25									
	17.3 [$p = 0.1250$]	$p = 0.0011$	15.0	$p = 0.1115$	14.0	$p = 0.0974$	13.1	$p = 0.5281$	12.7 [p > 0.1]

Notes: This table presents mean quantities per treatment using different subsets of market periods, and exact p-values from one-sided Wilcoxon rank sum tests and two-sided Wilcoxon signed rank tests. The rank sum test tests the inequalities between median prices in treatments. The Wilcoxon signed rank test tests the equality of observed mean quantities and predicted quantity in treatments No tax and Automatic (in squared brackets below the observed price). Tests are conducted using market level means across relevant periods.

Table A.9
Robustness of treatment effects on median price.

	No tax Predicted [158,162]	<	Seller only	<	Seller + BuyerC	<	Seller + Buyer	<	Automatic Predicted [178,182]
11–25									
	159.8 [p > 0.1]	p = 0.0119	168.3	p = 0.1407	172.1	p = 0.0400	178.0	p = 0.3918	178.3 [p > 0.1]
1–25									
	158.3 [p > 0.1]	p = 0.0076	168.3	p = 0.5260	170.0	p = 0.0162	176.7	0.1970	177.7 [p > 0.1]
9–25									
	159.7 [p > 0.1]	p = 0.0119	168.3	p = 0.1677	171.7	p = 0.0335	177.8	p = 0.5433	178.2 [p > 0.1]
10–25									
	160.0 [p > 0.1]	p = 0.0152	168.3	p = 0.1613	171.8	p = 0.0400	178.0	p = 0.5433	178.2 [p > 0.1]
12–25									
	159.8 [p > 0.1]	p = 0.0119	168.7	p = 0.1613	172.2	p = 0.0476	178.0	p = 0.2835	178.7 [p > 0.1]
13–25									
	159.8 [p > 0.1]	p = 0.0119	168.7	p = 0.1613	172.3	p = 0.0574	178.1	p = 0.3074	179.0 [p > 0.1]

Notes: This table presents averages of median market prices per treatment using different subsets of market periods, and exact p-values from one-sided Wilcoxon rank sum tests and Wilcoxon signed rank tests. The rank sum test tests the inequalities between median prices in treatments. The Wilcoxon signed rank test tests the equality of observed prices and midpoint of predicted prices in treatments No tax and Automatic (in squared brackets below the observed price). Tests are conducted using market level medians across relevant periods.

Table A.10
Wilcoxon rank sum and Kolmogorov-Smirnov test results for incidence, periods 11-25.

		Automatic exact p-value	p-value	Seller only exact p-value	p-value	Seller + Buyer exact p-value	p-value
Seller only	WRS-test	0.3095	0.2623				
	KS-test	0.1429	0.1389				
Seller + Buyer	WRS-test	0.6991	0.6310	0.3095	0.2623		
	KS-test	0.9307	0.8928	0.1429	0.1389		
Seller + BuyerC	WRS-test	0.2403	0.2002	0.1797	0.1495	0.2403	0.2002
	KS-test	0.4740	0.4413	0.1429	0.1389	0.4740	0.4413

Notes: This table reports the results of two-sided Wilcoxon rank sum tests and Kolmorov-Smirnov tests comparing incidence of observed effective unit tax between treatments. Reported are the exact and approximate p-values. The unit of observation is the mean incidence of observed effective per unit tax from main periods (11–25) in a given market. Observed effective unit tax is defined as the total of taxes and fines paid divided by the number of trades in a given market. Share of the tax borne by buyers is calculated as the difference in mean price in a market minus the mean price in the No tax treatment divided by the observed effective unit tax.

Table A.11
Wilcoxon rank sum and Kolmogorov-Smirnov test results for incidence, periods 1-25.

		Automatic exact p-value	p-value	Seller only exact p-value	p-value	Seller + Buyer exact p-value	p-value
Seller only	WRS-test	0.0931	0.0782				
	KS-test	0.1429	0.1389				
Seller + Buyer	WRS-test	0.3095	0.2623	0.0649	0.0547		
	KS-test	0.9307	0.8928	0.0260	0.0310		
Seller + BuyerC	WRS-test	0.0152	0.0163	0.0260	0.0250	0.2403	0.2002
	KS-test	0.1429	0.1389	0.0260	0.0310	0.4740	0.4413

Notes: This table reports the results of two-sided Wilcoxon rank sum tests and Kolmorov-Smirnov tests comparing incidence of observed effective unit tax between treatments. Reported are the exact and approximate p-values. The unit of observation is the mean incidence of observed effective per unit tax from periods 1–25 in a given market. Observed effective unit tax is defined as the total of taxes and fines paid divided by the number of trades in a given market. Share of the tax borne by buyers is calculated as the difference in mean price in a market minus the mean price in the No tax treatment divided by the observed effective unit tax.

Table A.12
Sellers' average trading prices, number of trades and risk aversion.

	Periods 11–25		Periods 1–25	
	Avg. price	Quantity	Avg. price	Quantity
AUTOMATIC	17.9***	-1.34**	15.4***	-1.29***
	(1.46)	(0.39)	(3.02)	(0.25)
SELLER ONLY	8.68*	-0.23	9.25*	-0.23
	(3.94)	(0.33)	(4.13)	(0.34)
SELLER + BUYER	15.2**	-0.74**	12.9*	-0.86**
	(4.19)	(0.18)	(5.14)	(0.21)
SELLER + BUYERC	8.24	-0.95***	4.04	-0.97***
	(5.09)	(0.081)	(5.57)	(0.060)
Risk aversion	-0.44	-0.026	-0.47	-0.037
	(0.51)	(0.024)	(0.58)	(0.018)
AUTOMATIC×Risk aversion	0.12	0.080	0.57	0.087
	(0.51)	(0.065)	(0.79)	(0.048)
SELLER ONLY×Risk aversion	0.20	-0.060	0.30	-0.058
	(0.78)	(0.063)	(0.81)	(0.063)
SELLER + BUYER×Risk aversion	0.73	-0.027	0.76	0.0061
	(0.74)	(0.033)	(0.95)	(0.034)
SELLER + BUYERC×Risk aversion	1.09	0.052	1.41	0.064
	(0.96)	(0.044)	(0.99)	(0.033)
Constant	161.4***	3.62***	160.4***	3.68***
	(2.73)	(0.070)	(2.68)	(0.070)
Observations	2250	2250	3749	3750
R^2	0.503	0.206	0.302	0.179

Notes: Dependent variable Avg. price is the mean price of the seller's traded units, Quantity the number of trades and reporting rate the share of reported units in a given period. The regressions in columns 1 and 2 use data from periods 11–25 and from periods 1–25 in columns 3 and 4. The constant term captures the estimate for the NO TAX treatment. Risk preferences were elicited in the post-experimental questionnaire by directly asking subjects to assess the following question taken from the German Socio-Economic Panel: "How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'not at all willing to take risks' and the value 10 means: 'very willing to take risks'." Risk aversion is the risk attitude measurement re-scaled (10 - Risk attitude score). Standard errors clustered at the market and subject level are in parentheses.

Table A.13
Sellers' average trading prices, number of trades, reporting rates and risk aversion in treatments with evasion opportunities, periods 11–25.

	(1) Avg. price	(2) Trades	(3) Reporting rate
SELLER + BUYER	6.54	-0.51	0.63**
	(3.51)	(0.44)	(0.16)
SELLER + BUYERC	-0.44	-0.71	0.24
	(4.38)	(0.38)	(0.17)
Risk aversion	-0.24	-0.087	0.024
	(0.39)	(0.079)	(0.044)
SELLER + BUYER×Risk aversion	0.53	0.033	-0.015
	(0.58)	(0.075)	(0.047)
SELLER + BUYERC×Risk aversion	0.89	0.11	0.022
	(0.73)	(0.076)	(0.046)
Constant	170.1***	3.39***	0.20
	(2.11)	(0.35)	(0.15)
Observations	1350	1350	1350
R^2	0.229	0.078	0.346

Notes: Dependent variable Avg. price is the mean price of the seller's traded units, Trades the number of trades and reporting rate the share of reported units in a given market-period. The regressions use data from periods 11–25. The constant term captures the estimate for the SELLER ONLY treatment. Risk preferences were elicited in the post-experimental questionnaire by directly asking subjects to assess the following question taken from the German Socio-Economic Panel: "How do you see yourself: are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 0 means: 'not at all willing to take risks' and the value 10 means: 'very willing to take risks'." Risk aversion is the risk attitude measurement re-scaled (10 - Risk attitude score). Standard errors clustered at the market and subject level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14
Effect of treatment on seller reporting rate.

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	0.32***	0.38***	0.33***	0.47*	0.40***
	(0.031)	(0.026)	(0.044)	(0.26)	(0.13)
Seller + Buyer	0.51***	0.36***	0.51***	0.43	0.36**
	(0.041)	(0.042)	(0.062)	(0.26)	(0.14)
Seller + BuyerC	0.32***	0.21***	0.36***	0.36	0.20
	(0.057)	(0.063)	(0.077)	(0.28)	(0.18)
Period		-0.004***			
		(0.001)			
Seller + Buyer * Period		0.011***			
		(0.002)			
Seller + BuyerC * Period		0.008***			
		(0.002)			
Tax morale (bin.)			-0.010		
			(0.12)		
Seller + Buyer * Tax morale (bin.)			0.0057		
			(0.14)		
Seller + BuyerC * Tax morale (bin.)			0.13		
			(0.14)		
Risk attitude				-0.026	
				(0.040)	
Seller + Buyer * Risk att.				0.013	
				(0.041)	
Seller + BuyerC * Risk att.				-0.011	
				(0.046)	
Generalized trust					-0.014
					0.021
Seller + Buyer * Gen. trust					0.025
					(0.022)
Seller + BuyerC * Gen. trust					0.020
					(0.031)
R^2	0.26	0.27	0.27	0.28	0.27
Observations	2249	2249	2249	2249	2249

Notes: Dependent variable Reporting rate is the ratio of the number of units reported to the total number of units sold by a seller in a given market period. Used are all periods from 1 to 25. One individual-period observation is excluded due to zero units sold. The constant term captures the estimate for the Seller Only treatment. Tax morale (bin.) is a binary variable that takes value 1 if the seller reports cheating on taxes never being acceptable. Robust standard errors clustered at the market level are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table A.15
Additional descriptive statistics, periods 11-25.

	No tax		Seller only		Seller + BuyerC		Seller + Buyer		Automatic
Division of revenue									
Total earnings	21,221		17,091		12,199		11,847		11,949
Earnings of sellers	10,363		8,741		4,938		5,309		5,578
Earnings of buyers	10,858		8,191		7,261		6,539		6372
Taxes collected	-		3000		6,487		7,493		7,707
Fines	-		473		1093		573		-
Efficiency									
Maximum efficiency	21,420								20280
Observed efficiency	21,221	≠ **	20,564	≠ *	19,779	≠	19,914	≠	19,656
Relative efficiency	100%		96.9%		93.2%		93.8%		92.6%

Notes: Earnings, taxes collected and fines are averages per session across relevant periods and expressed in ECU. Maximum efficiency is the maximum sum of seller and buyer surplus. Observed efficiency of a treatment is the mean total sum of earnings, collected taxes and fines. Relative efficiency is defined as the efficiency of a given treatment divided by the efficiency in the No tax treatment. The greater than and smaller than signs specify the direction of a one-sided Wilcoxon rank sum test and unequal sign indicates a two-sided test. Tests are conducted using market level means across relevant periods. The stars designate conventional levels of statistical significance (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Observations from periods 11–25.

Table A.16
Additional descriptive statistics, periods 1-25.

	No tax		Seller only		Seller + BuyerC		Seller + Buyer		Automatic
Division of revenue									
Total earnings	35,302		28,311		19,545		19,128		19,790
Earnings of sellers	16,782		14,271		6,143		7140		8,627
Earnings of buyers	18,520		13,694		13,402		11,988		11,163
Taxes collected	-		5,320		11,047		12,613		13,260
Fines	-		793		2,300		1,433		-
Efficiency									
Maximum efficiency	35,700								33,800
Observed efficiency	35,302	≠ **	34,424	≠ **	32,892	≠	33,174	≠	33,050
Relative efficiency	100%		97.5%		93.2%		94.0%		93.6%

Notes: Earnings, taxes collected and fines are averages per session across relevant periods and expressed in ECU. Maximum efficiency is the maximum sum of seller and buyer surplus. Observed efficiency of a treatment is the total sum of earnings, collected taxes and fines. Relative efficiency is defined as the efficiency of a given treatment divided by the efficiency in the No tax treatment. The greater than and smaller than signs specify the direction of a one-sided Wilcoxon rank sum test and unequal sign indicates a two-sided test. Tests are conducted using market level means across relevant periods. The stars designate conventional levels of statistical significance (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). Observations from periods 1–25.

Appendix B. Additional figures

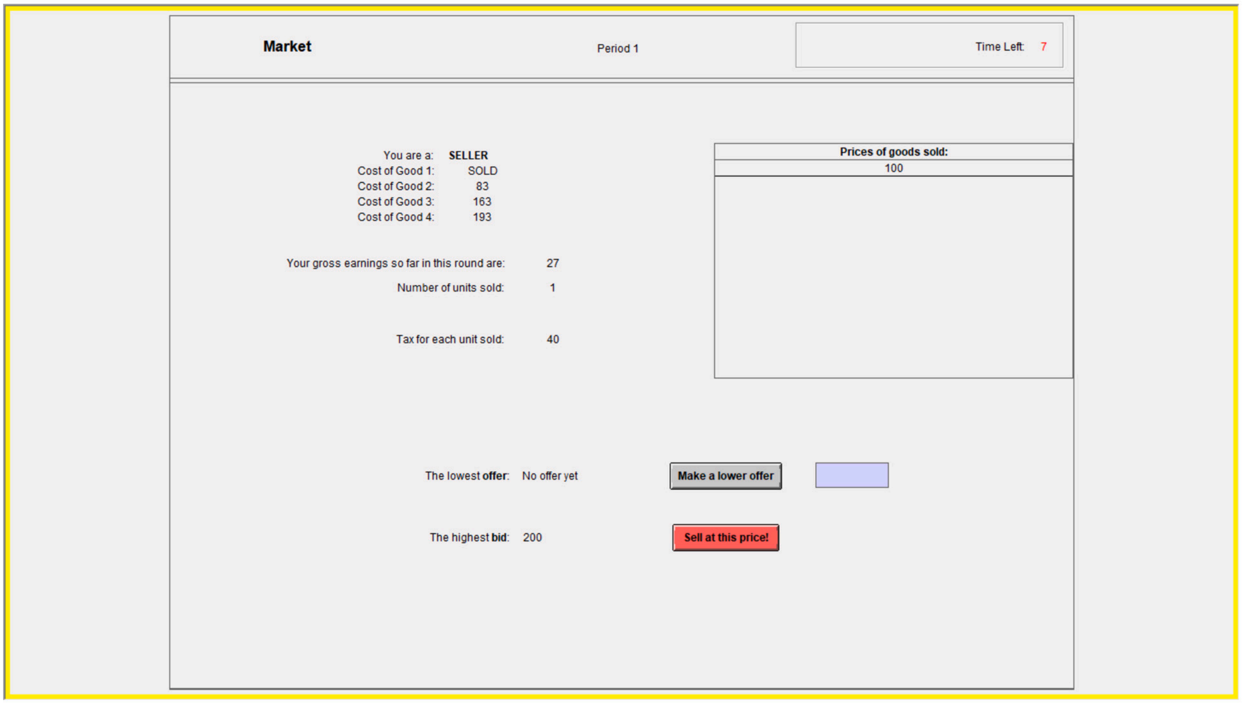

Fig. B.1. Seller's trading screen.

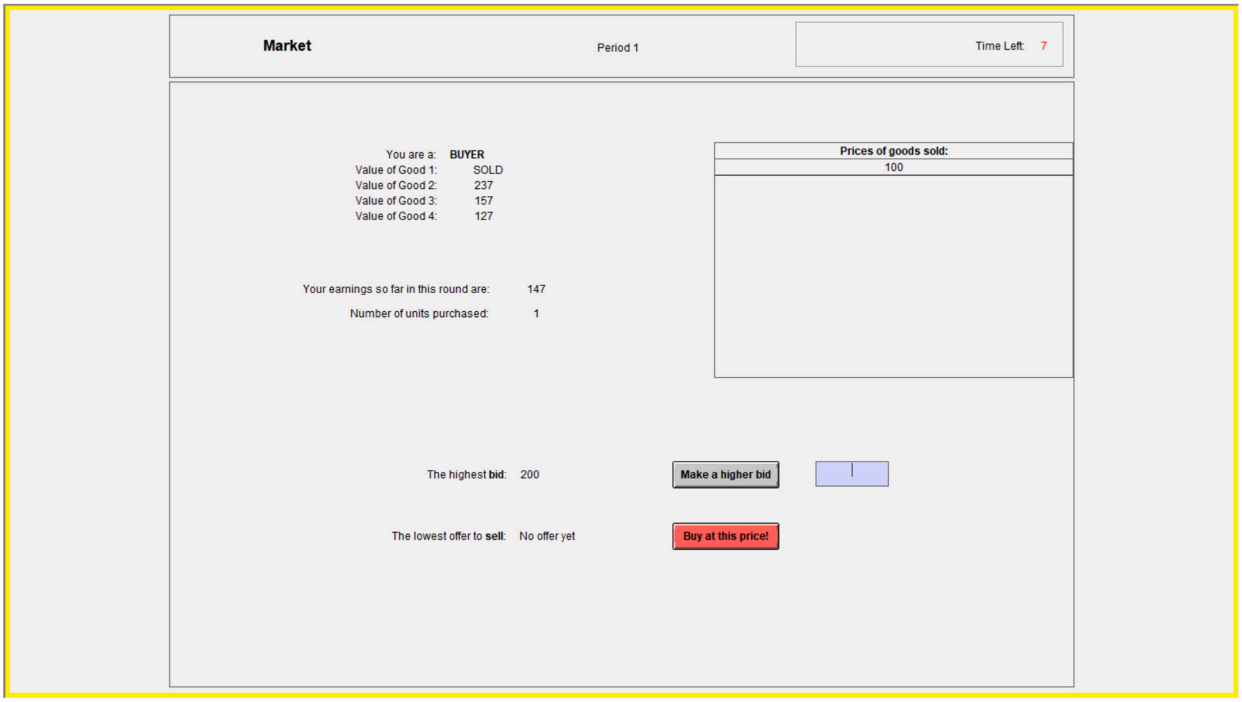

Fig. B.2. Buyer's trading screen.

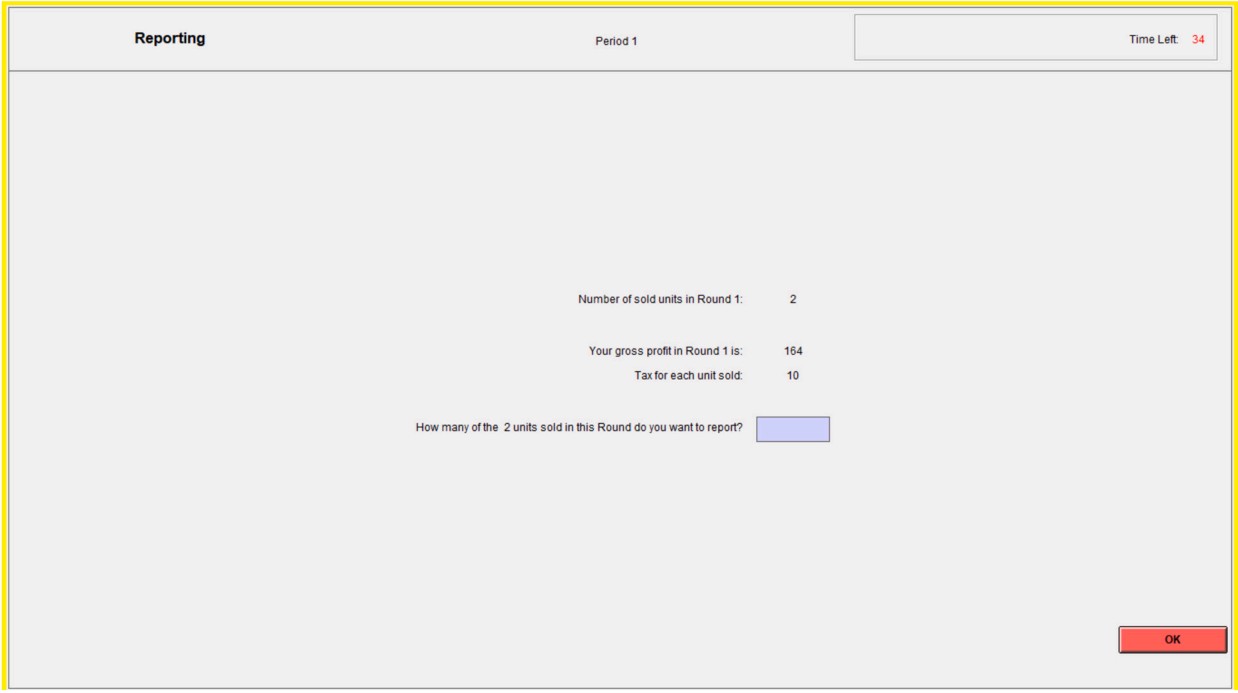

Fig. B.3. Seller's reporting screen.

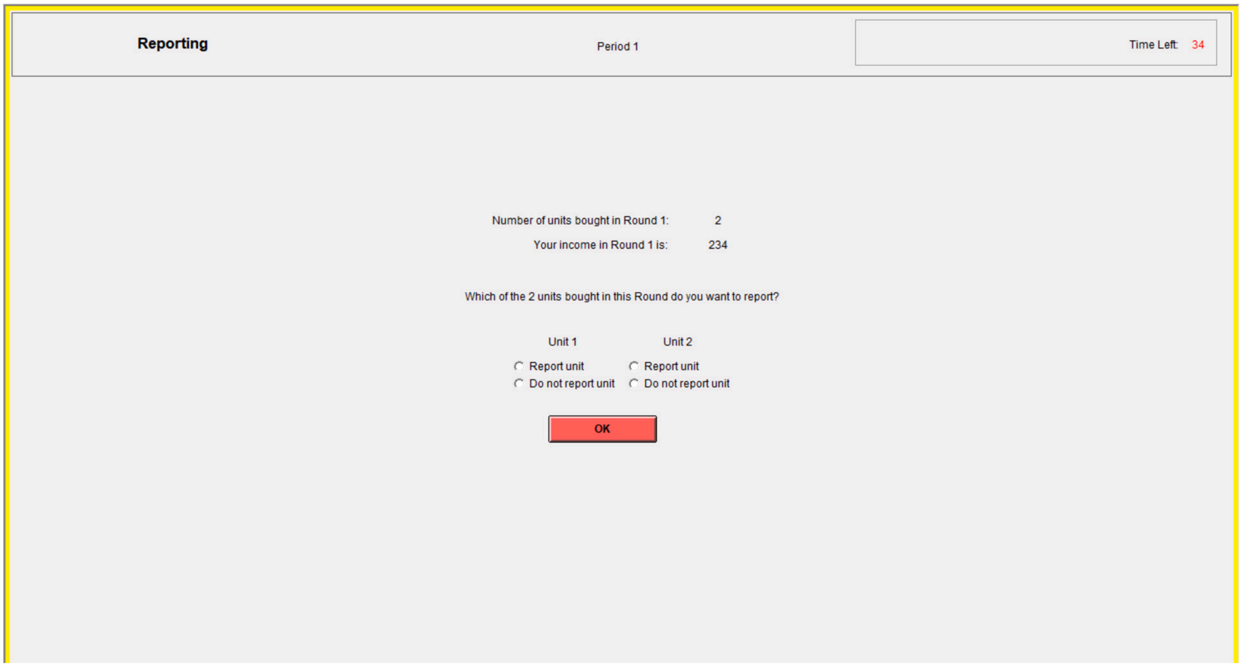

Fig. B.4. Buyer's reporting screen.

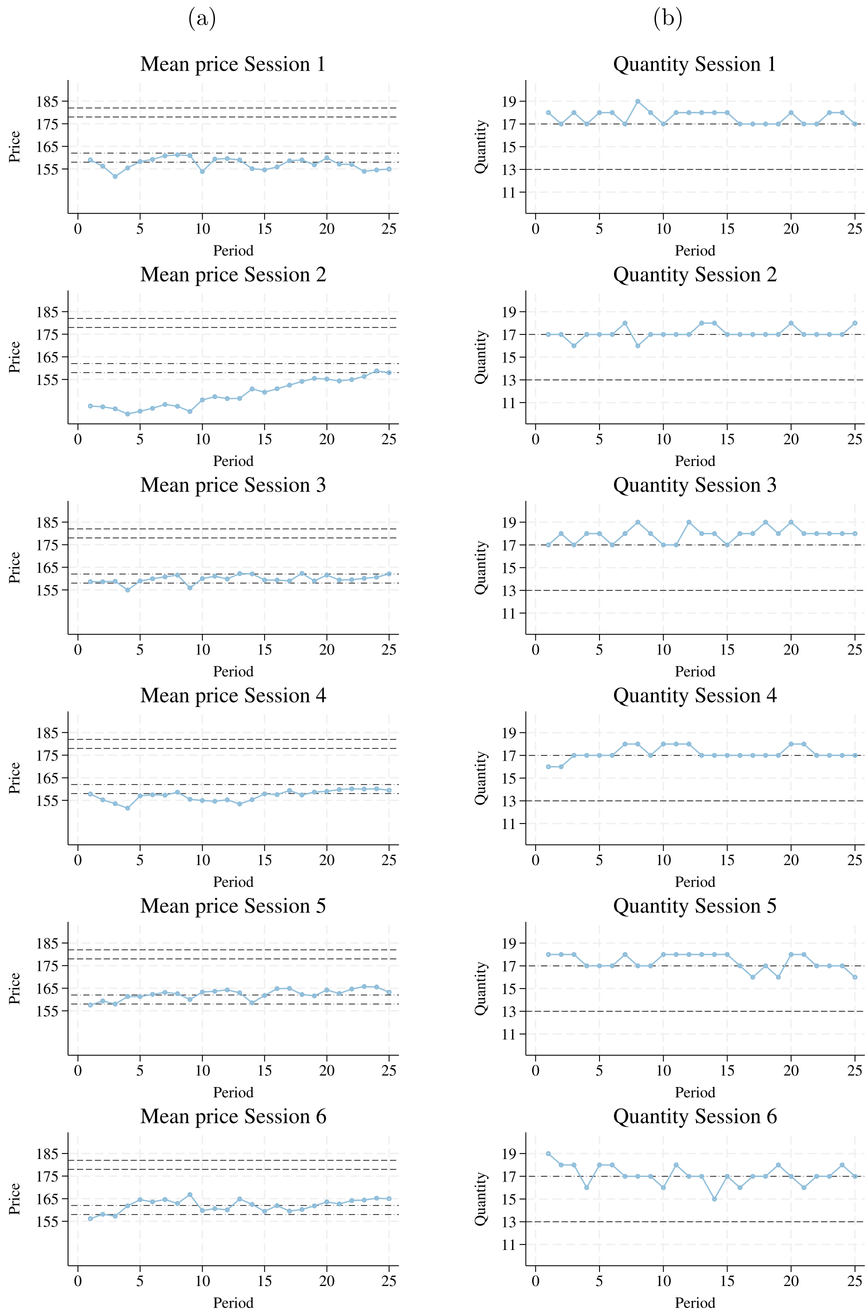

Fig. B.5. Evolution of mean prices and quantities by session in treatment No tax.

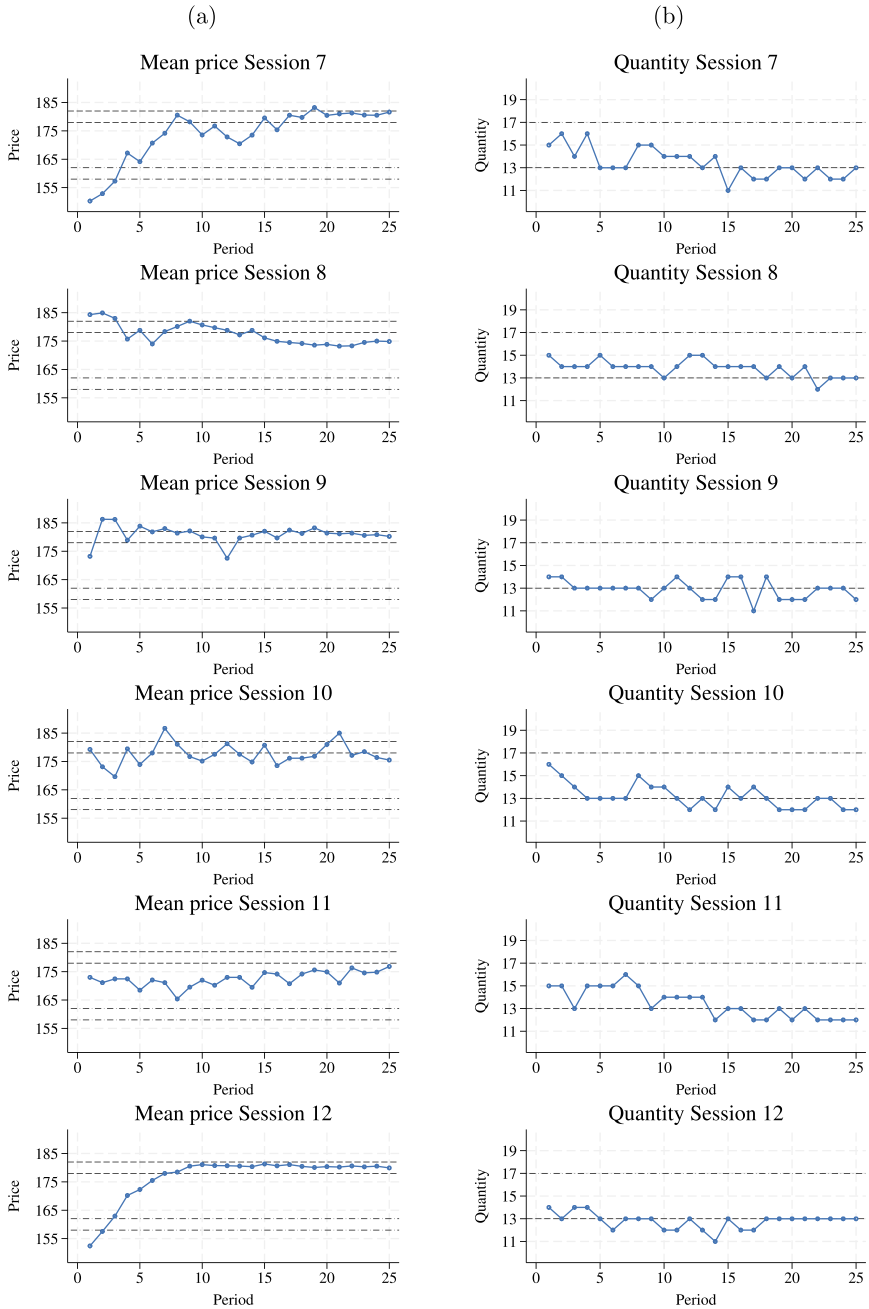

Fig. B.6. Evolution of mean prices and quantities by session in treatment AUTOMATIC.

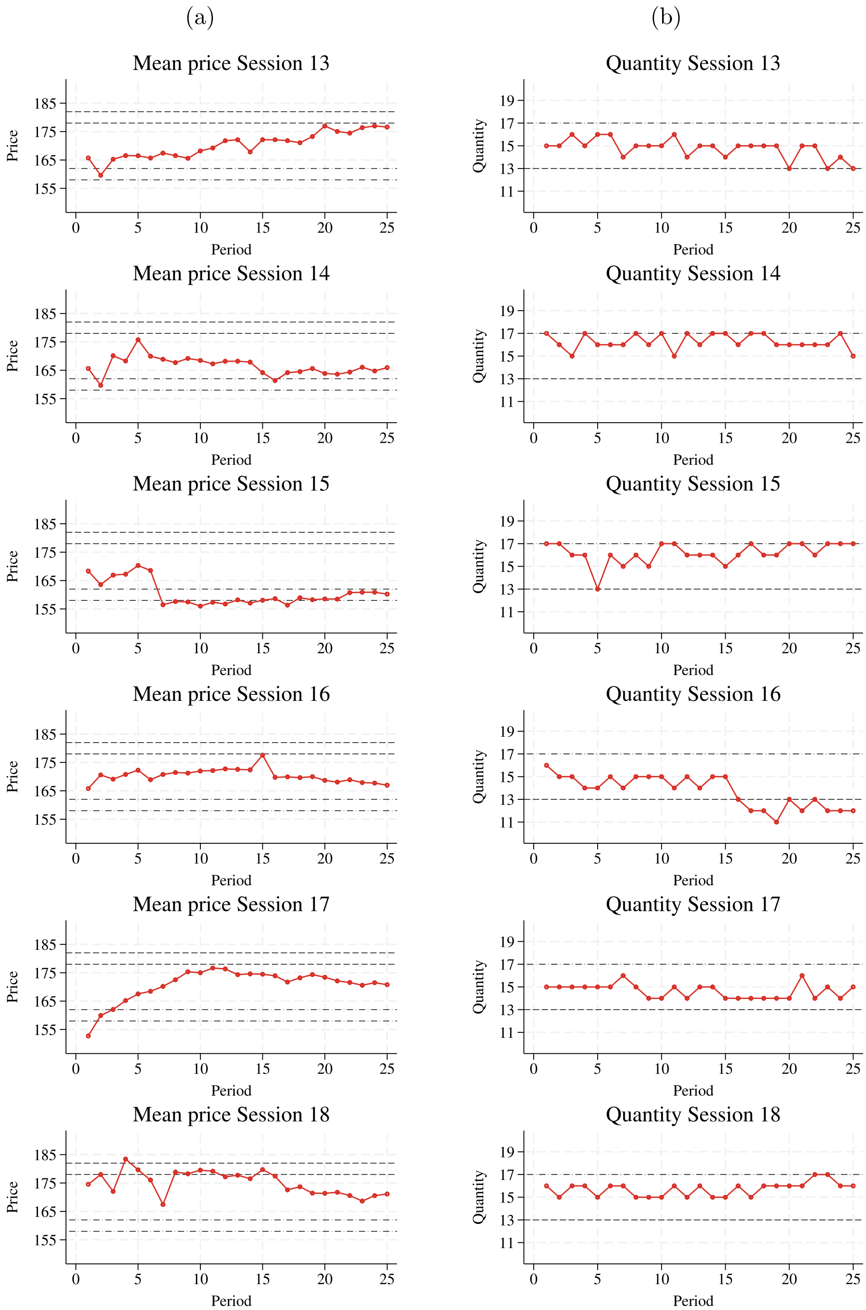

Fig. B.7. Evolution of mean prices and quantities by session in treatment SELLER ONLY.

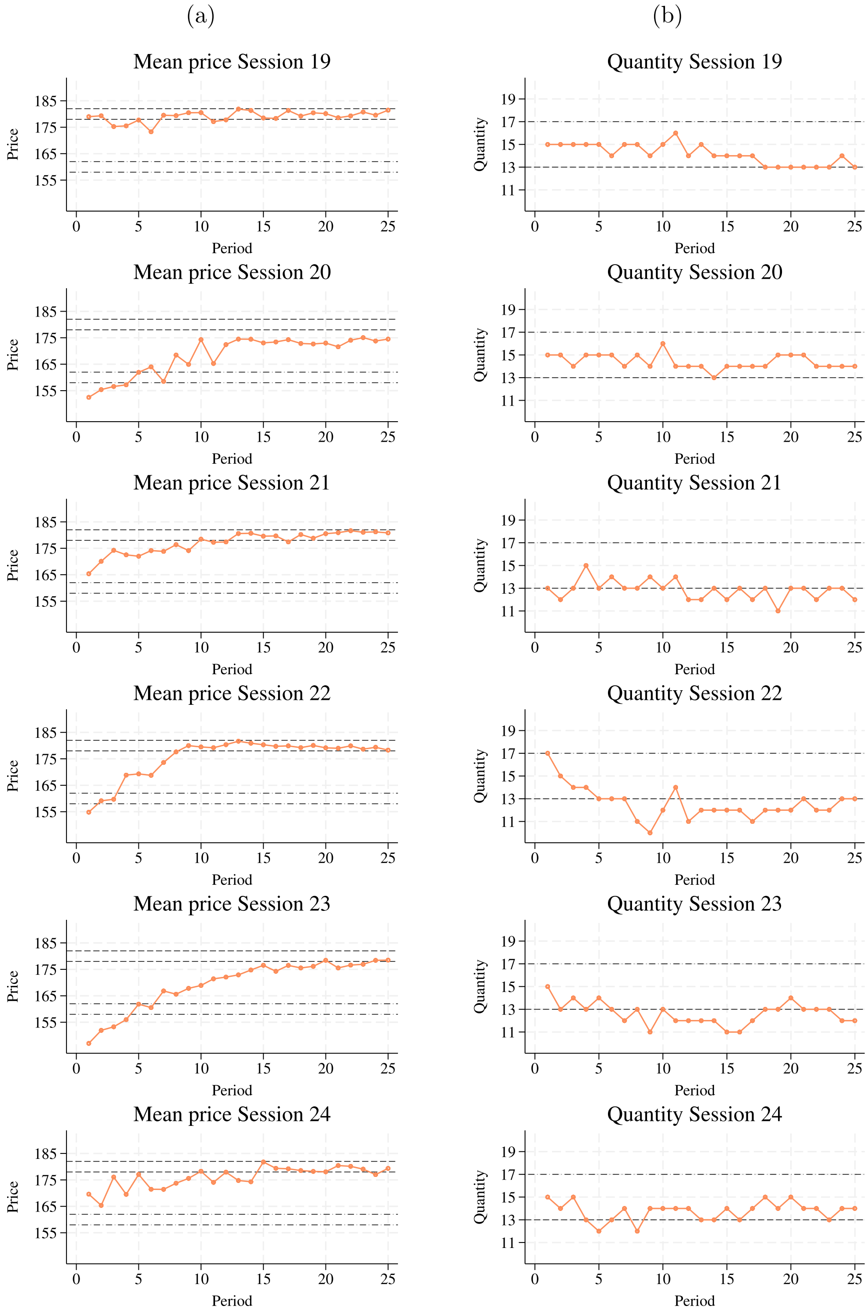

Fig. B.8. Evolution of mean prices and quantities by session in treatment Seller + Buyer.

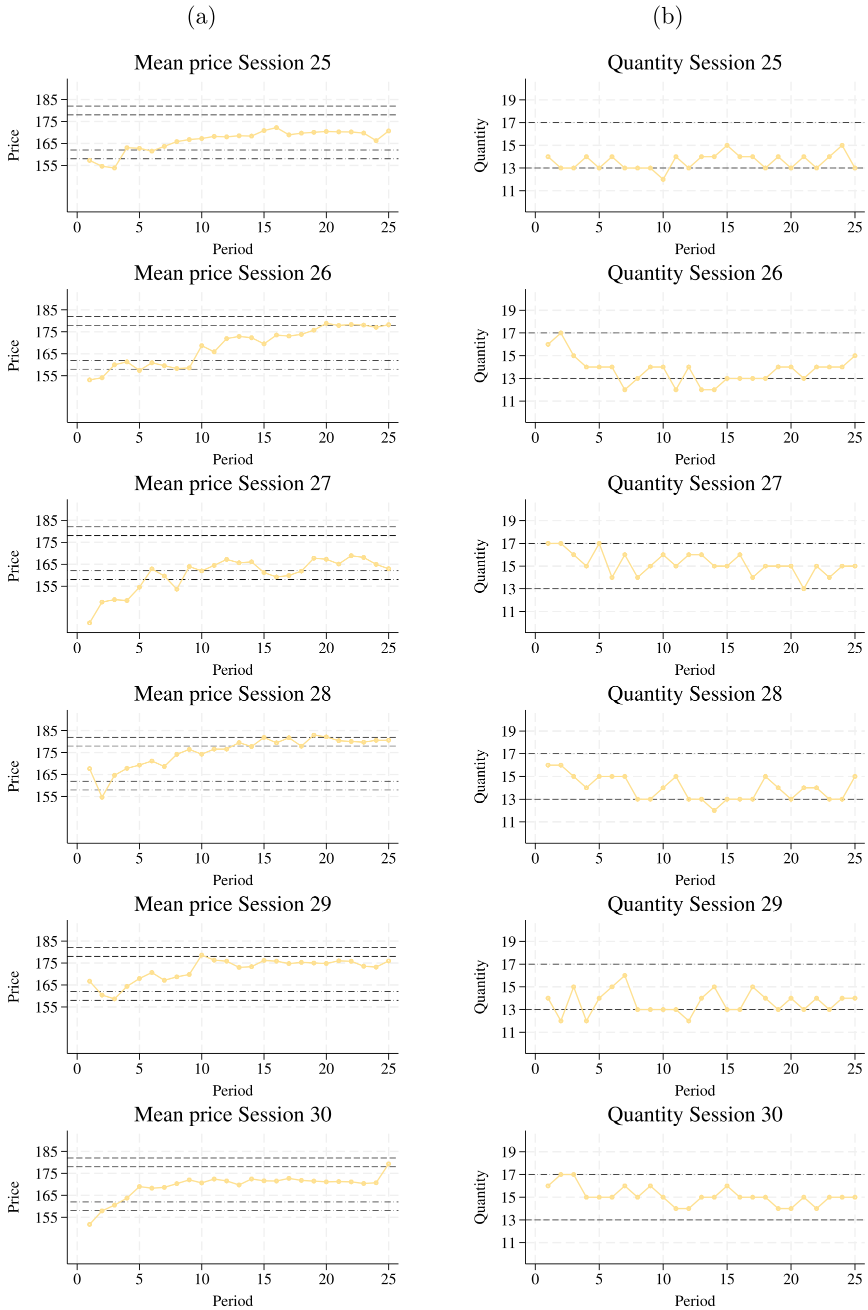

Fig. B.9. Evolution of mean prices and quantities by session in treatment SELLER + BUYERC.

(a)

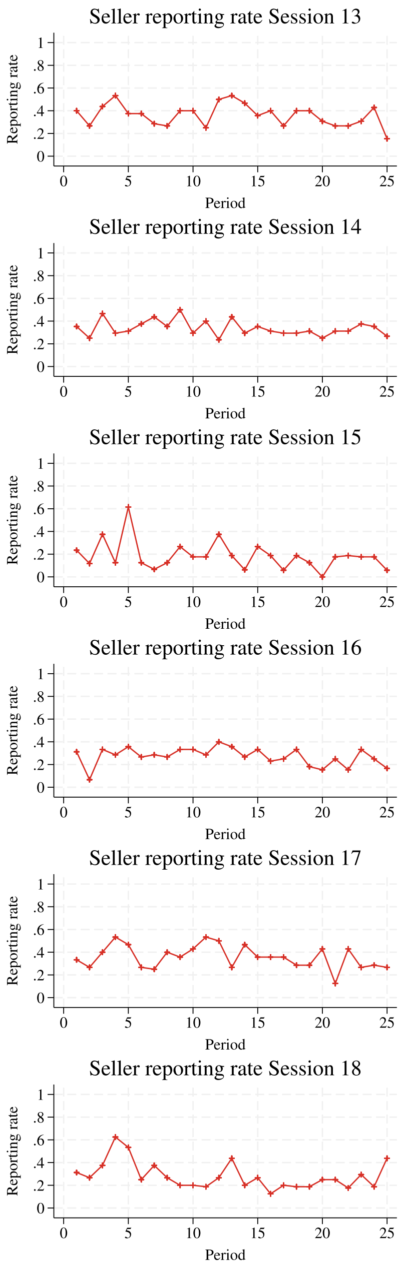

Fig. B.10. Evolution of seller reporting rates by session in treatment Seller only.

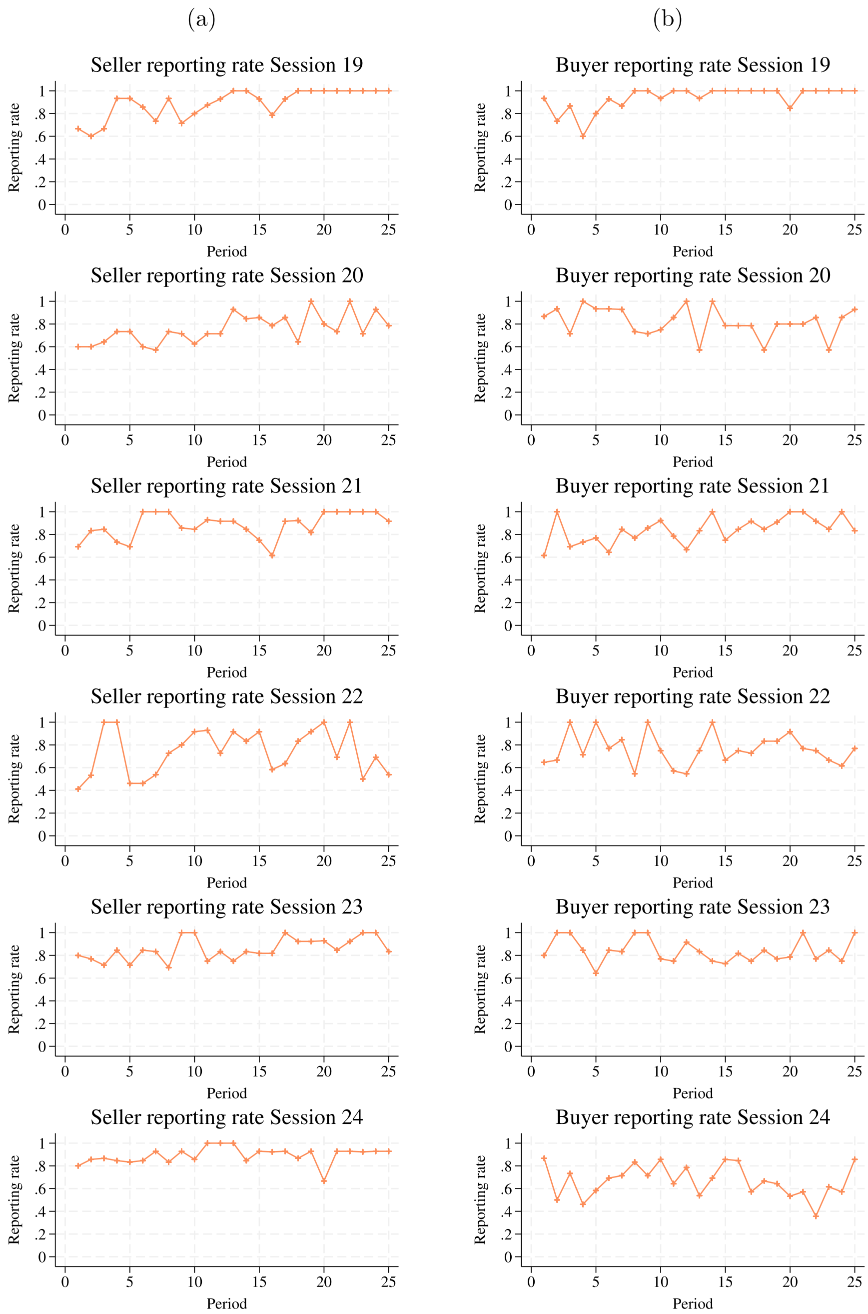

Fig. B.11. Evolution of seller and buyer reporting rates by session in treatment SELLER + BUYER.

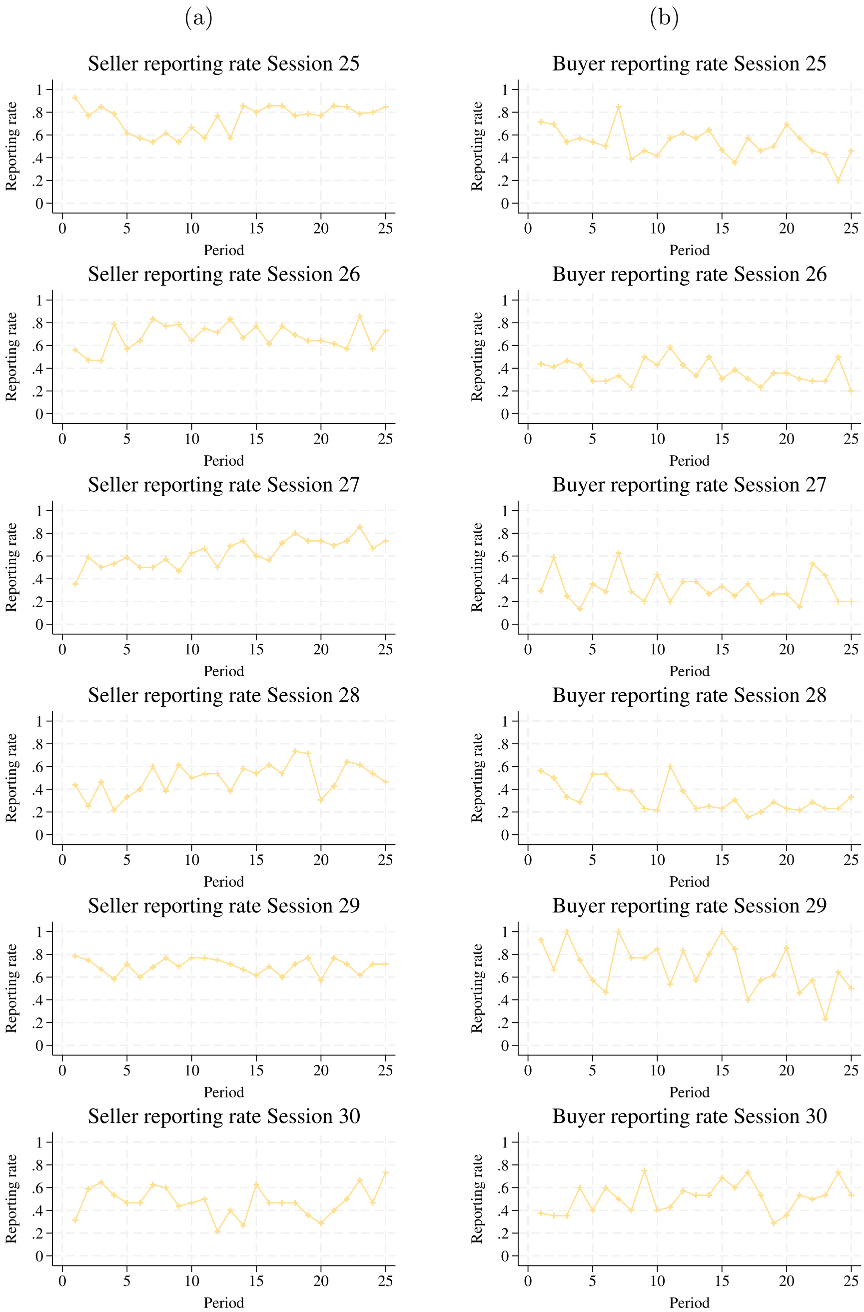

Fig. B.12. Evolution of seller and buyer reporting rates by session in treatment SELLER + BUYERC.

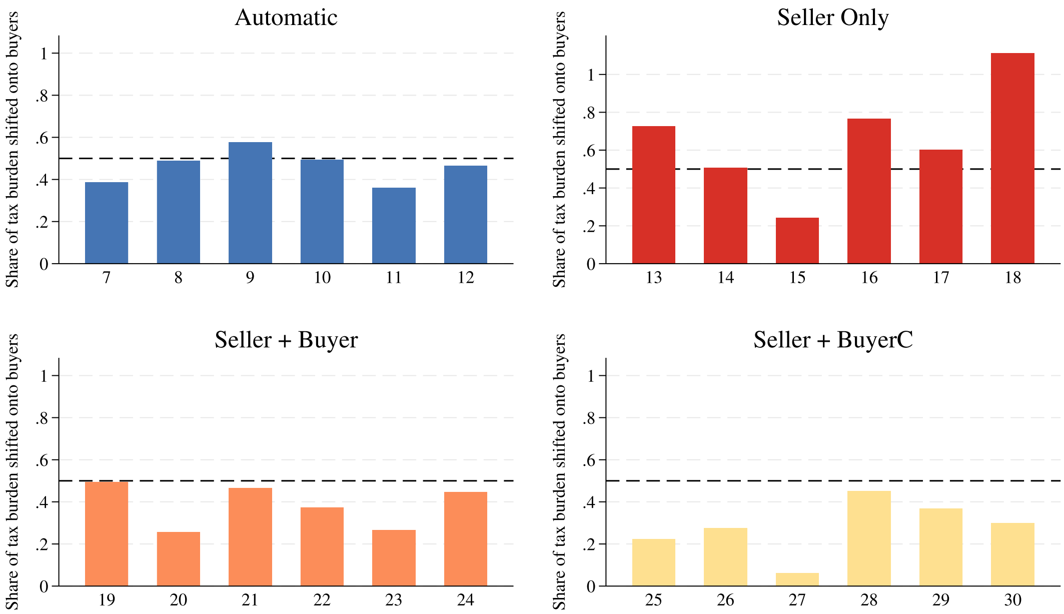

Fig. B.13. Incidence of observed effective tax, periods 1-25. The figure displays the incidence of the observed effective unit tax by market in treatments AUTOMATIC, SELLER ONLY, SELLER + BUYER and SELLER + BUYERC. Observed effective unit tax is defined as the total of taxes and fines paid divided by the number of trades in a given market. Share of the tax borne by buyers is calculated as the difference in mean price in a market minus the mean price in the NO TAX treatment divided by the observed effective unit tax. The figure shows incidences using data from periods 1-25 (see Fig. 5 for figures using data from periods 11-25).

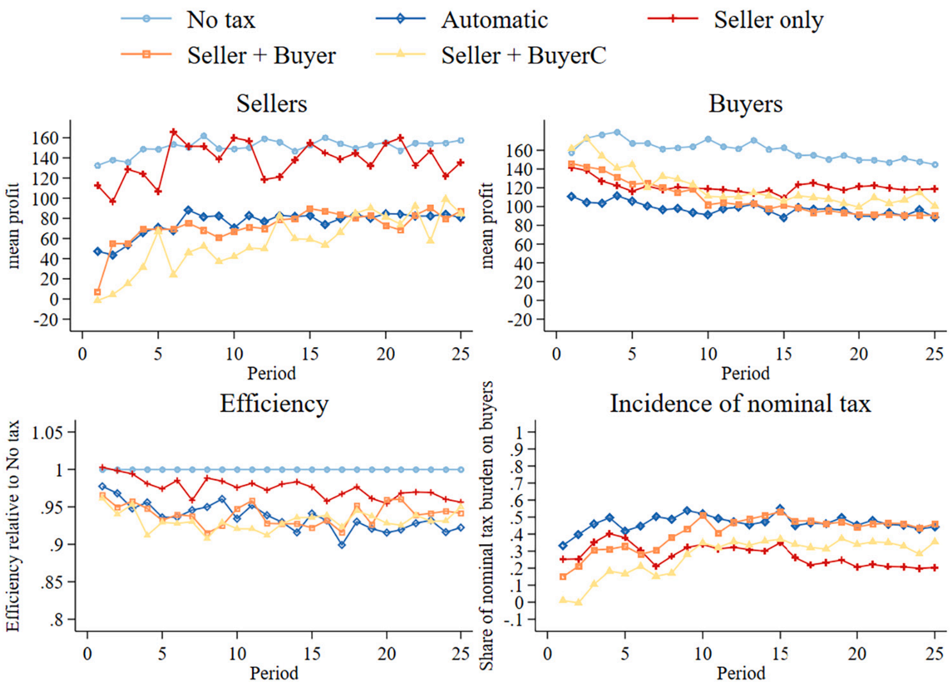

Fig. B.14. Evolution of profits, efficiency and incidence of nominal tax by treatment. The figure plots the evolution of mean profits of sellers and buyers, relative efficiency and incidence of nominal tax by treatment in periods 1-25.

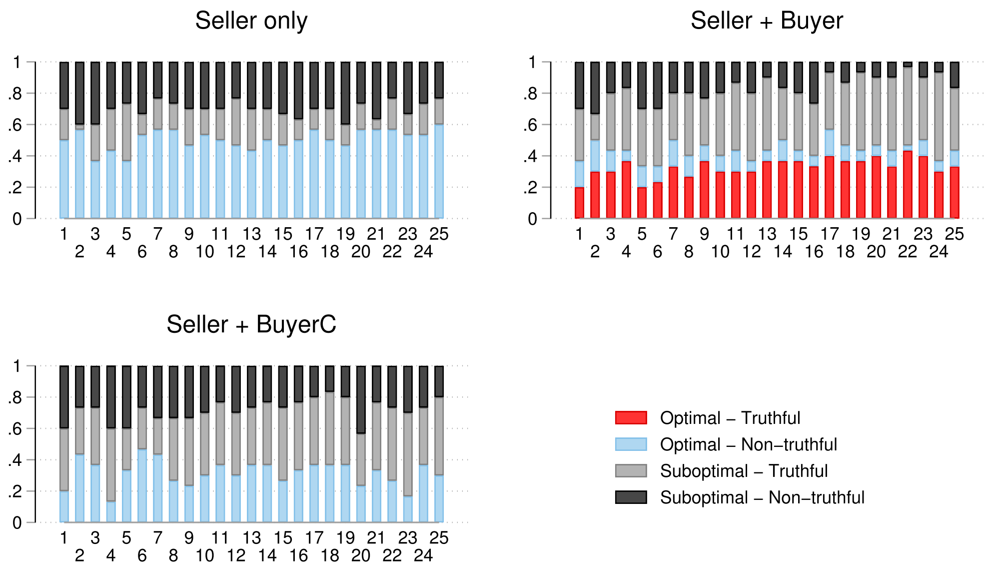

Fig. B.15. Evolution of sellers' report types by treatment. The figure plots the evolution of report types Optimal-Truthful, Optimal - Non-truhful, Suboptimal - Truhful, Suboptimal - Non-truthful across periods by treatments. Truthful refers to reporting all sold units. Optimality is defined using the expected effective audit rate. In SELLER ONLY the audit rate $\gamma_{SO} = 0.1$. In SELLER + BUYER and SELLER + BUYERC the effective audit probability depends on the number of units sold and reported, and the buyer reporting rate. These figures use the observed mean buyer reporting rates in periods 11-25 (81% in SELLER + BUYER and 43% in SELLER + BUYERC, yielding the probabilities reported in Table 3).

SELLER + BUYER

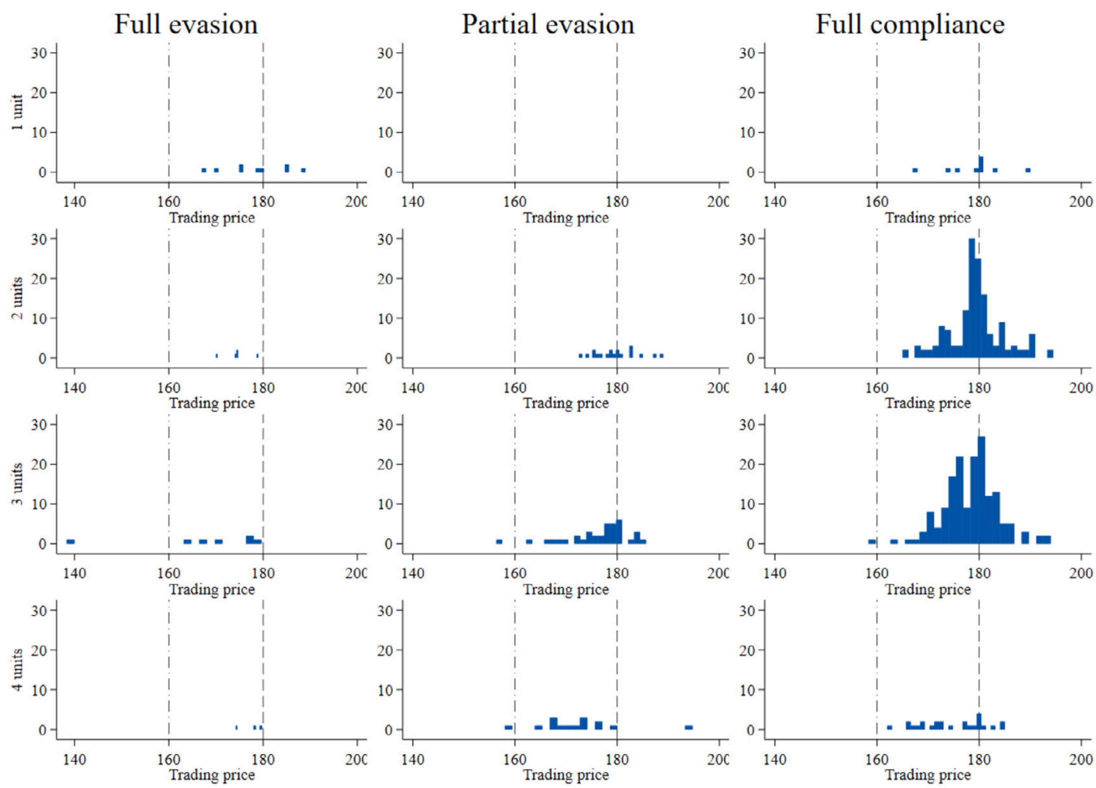

Fig. B.16. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in SELLER + BUYER treatment. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

SELLER + BUYERC

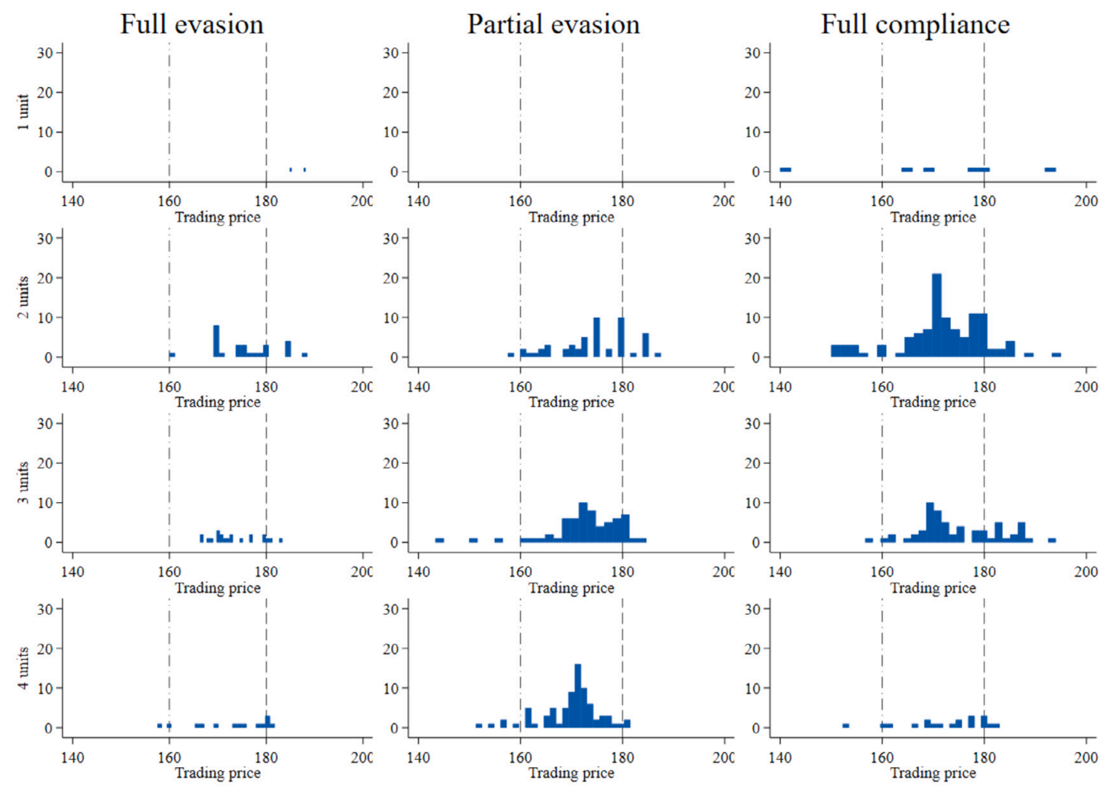

Fig. B.17. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in SELLER + BUYERC treatment. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

SELLER ONLY, Session 13

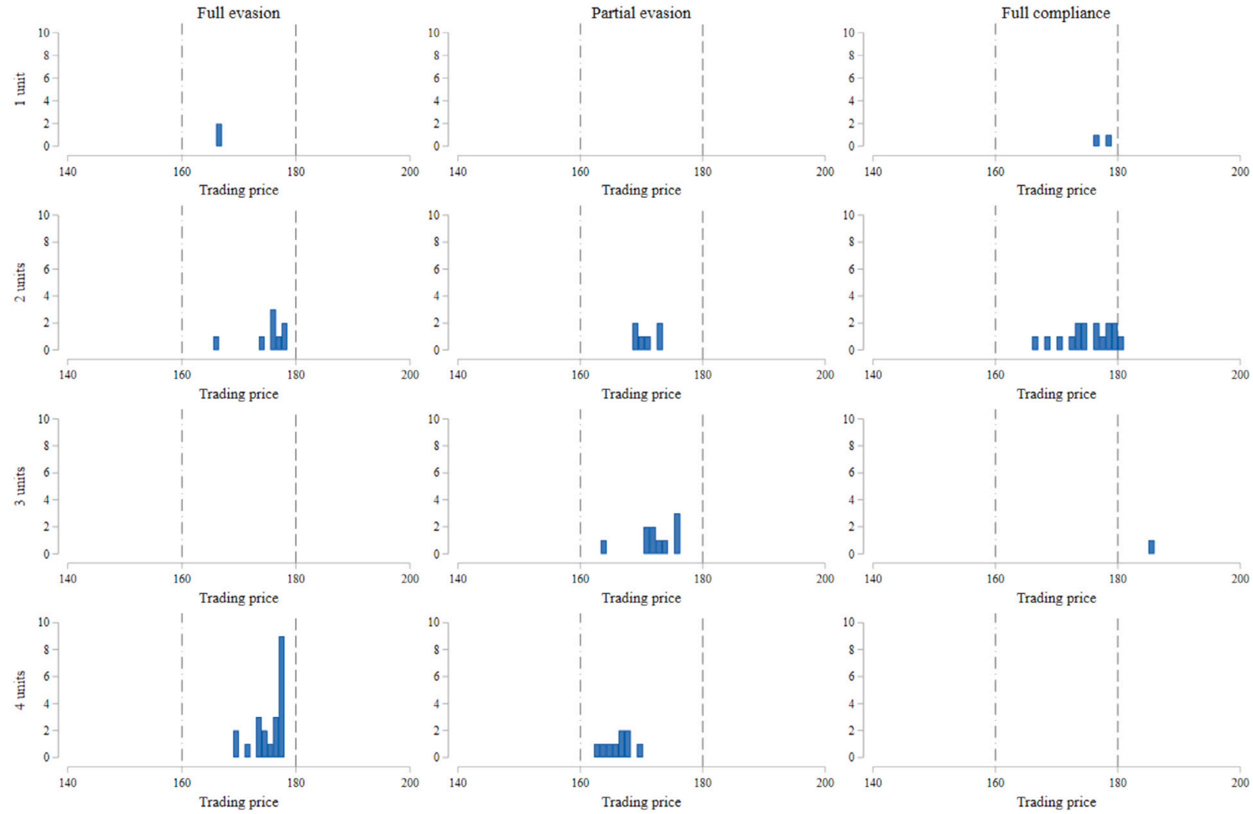

Fig. B.18. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in SELLER ONLY session 13. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Seller only, Session 14

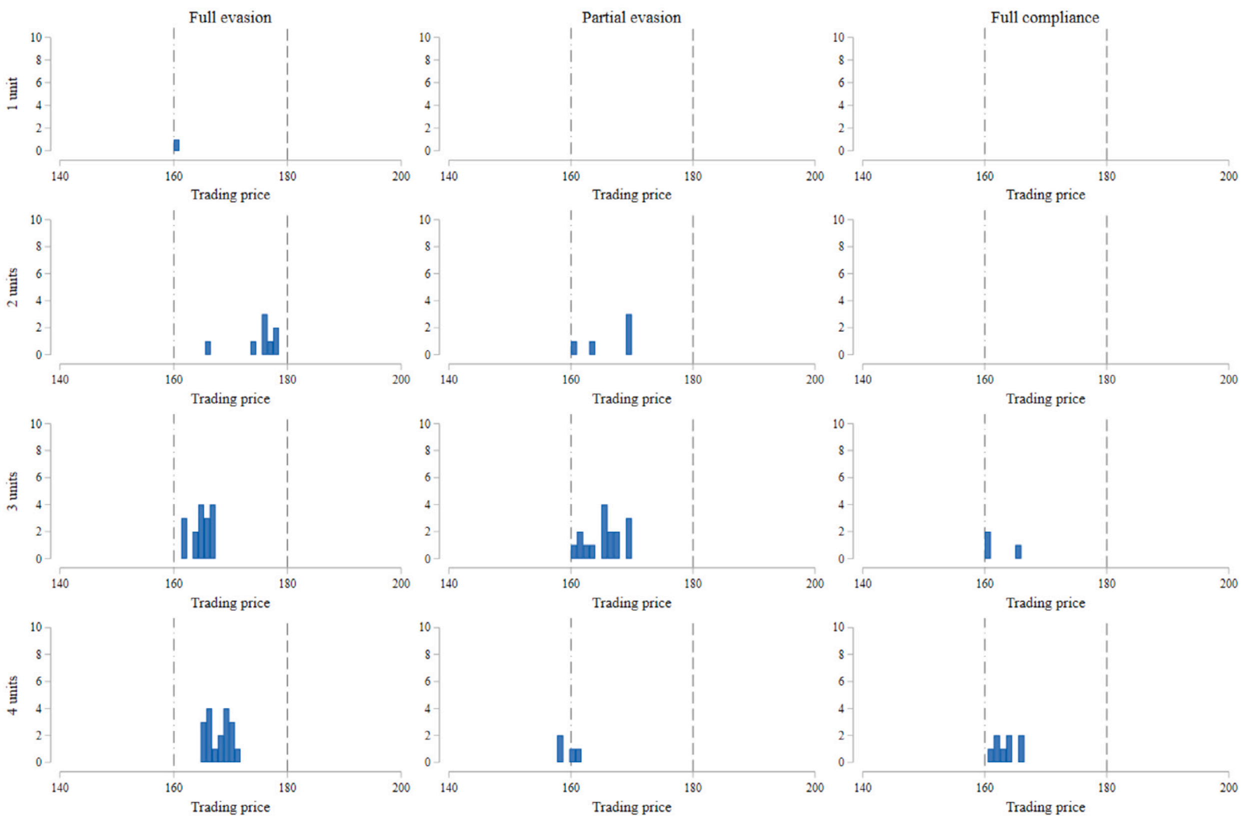

Fig. B.19. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in Seller only session 14. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Seller only, Session 15

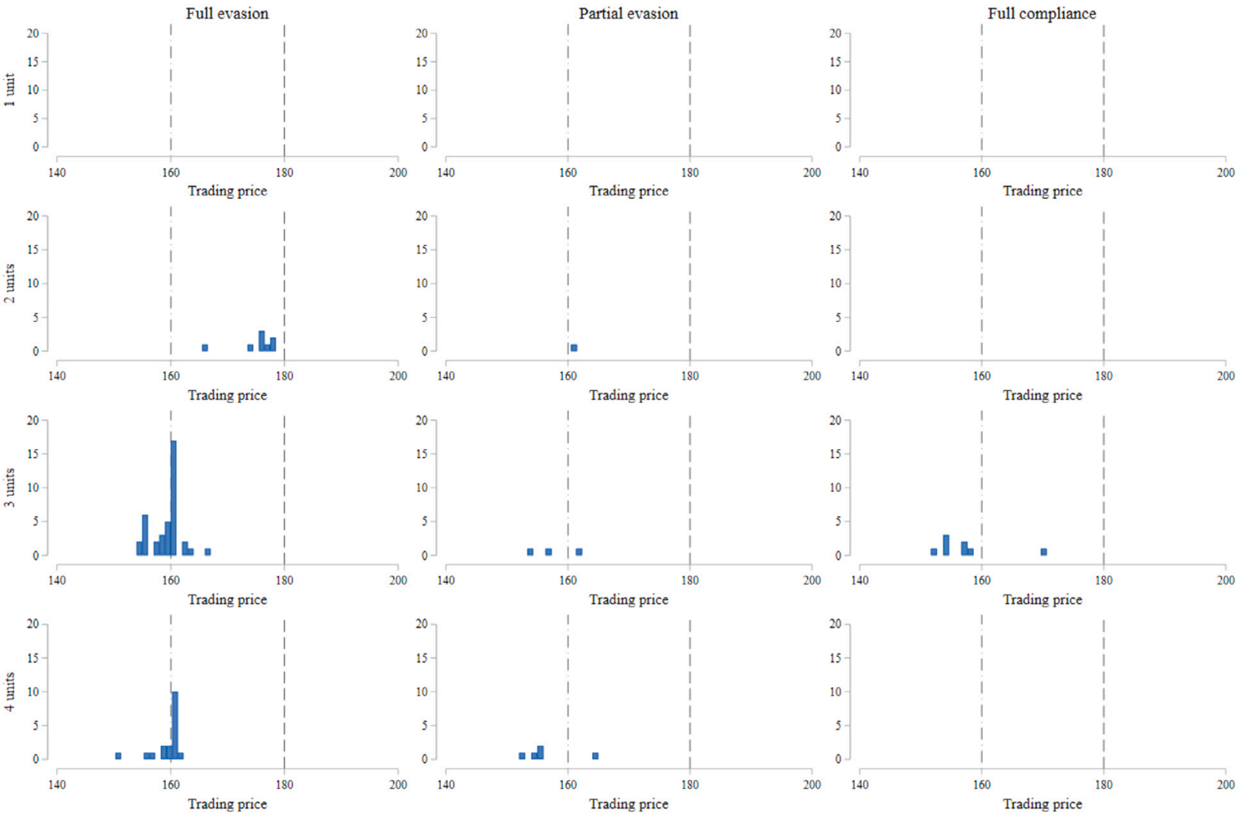

Fig. B.20. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in Seller only session 15. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Seller only, Session 16

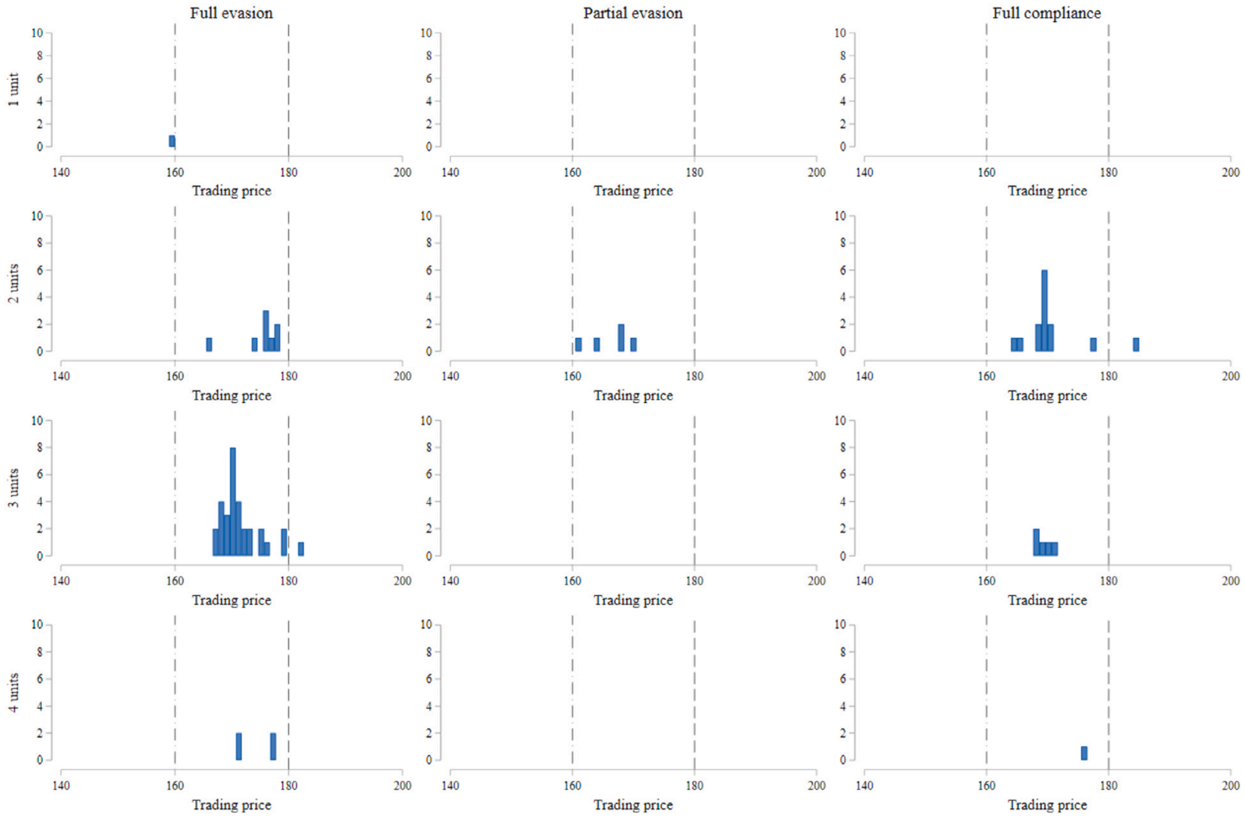

Fig. B.21. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in Seller only session 16. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Seller only, Session 17

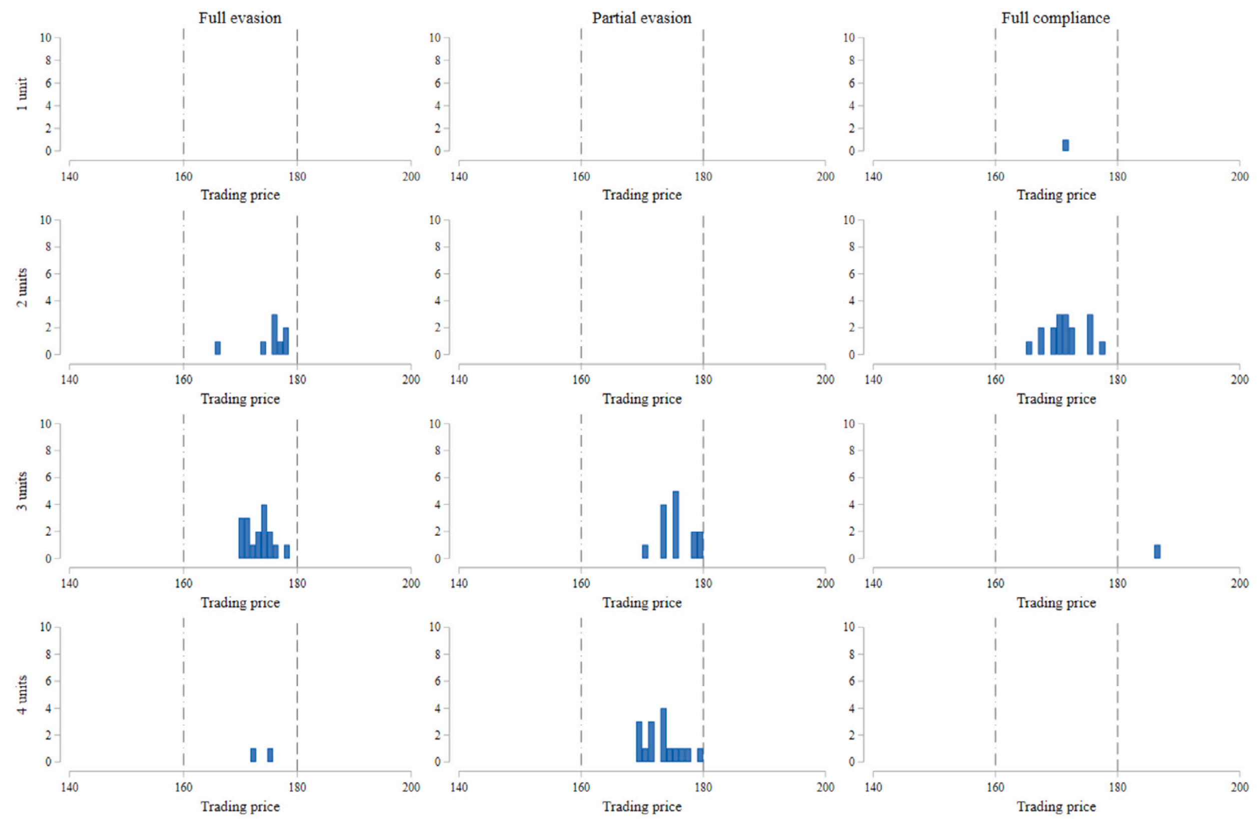

Fig. B.22. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in Seller only session 17. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Seller only, Session 18

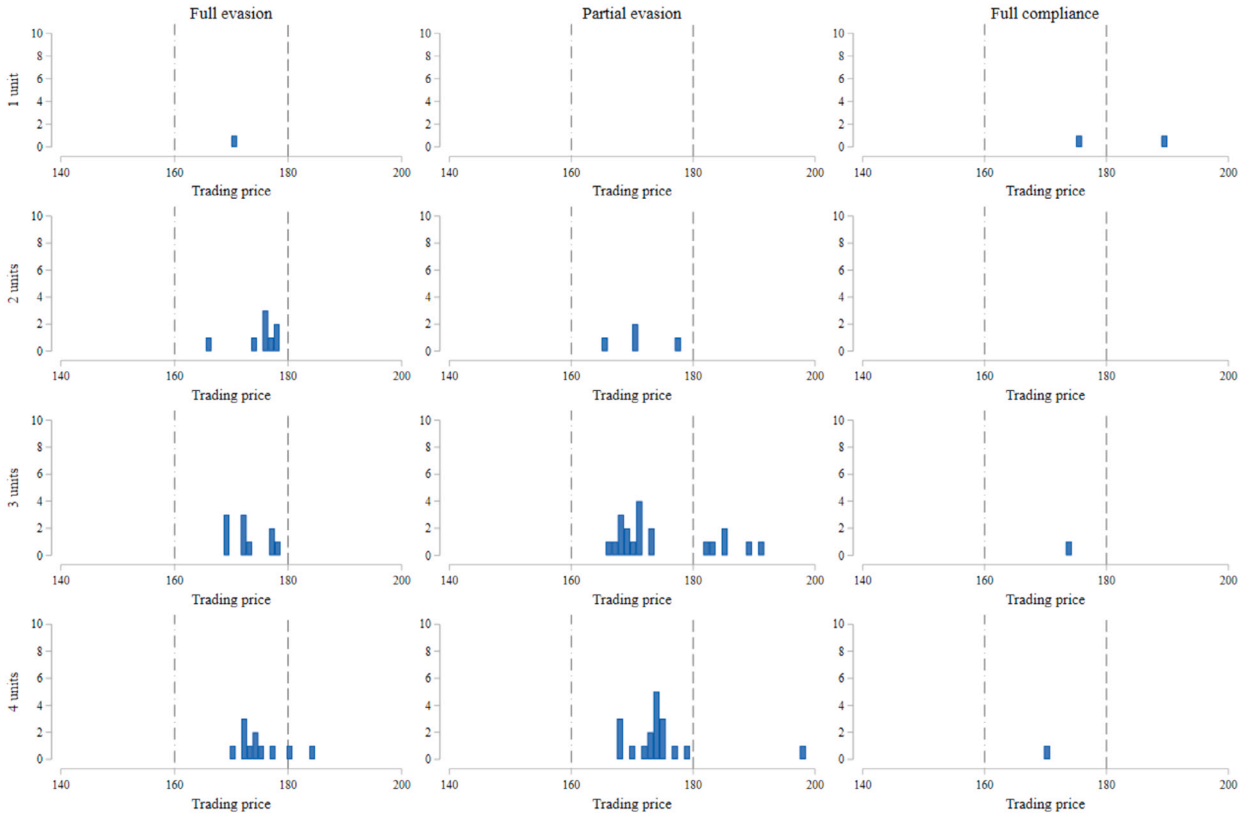

Fig. B.23. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in Seller only session 18. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

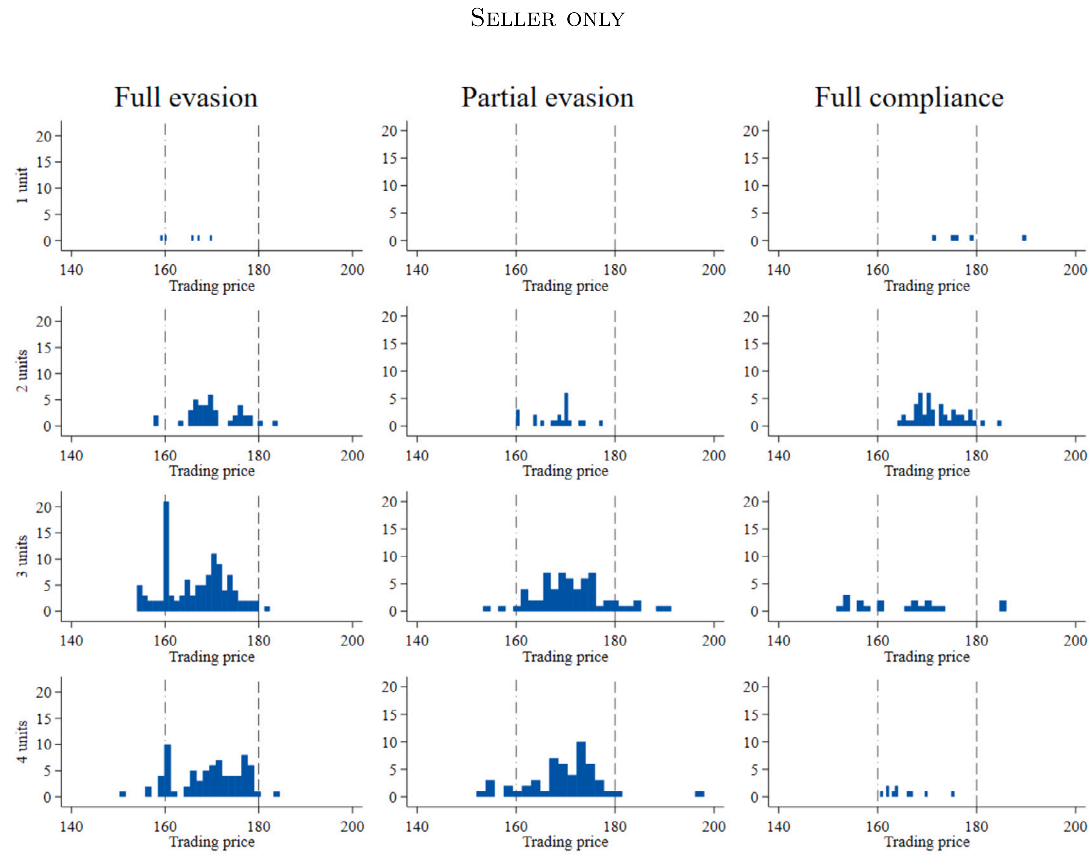

Fig. B.24. Trading prices by sold units and seller's reporting decision. Frequencies of average trading prices by number of units sold and seller's reporting decision using seller-period level data from periods 11–25 in Seller only treatment. Vertical lines indicate the predicted market clearing prices without taxes (160 ECU) and with taxes (180 ECU).

Appendix C. Derivation of predictions

In this section we formally derive the predictions of Section 4. Furthermore, by pinning down preferences of the traders, we are able to derive more precise predictions than the pre-registered predictions in Section 4. Throughout, we assume that sellers are self-interested and risk-neutral, and seek to maximize their monetary earnings.

C.1. Reporting behavior and expected tax liability

By not reporting some of the sold units in period t, $s_{i,t}$, seller i avoids having to pay the tax of 40 ECU for these units. At the same time, she faces a risk of being audited and having to pay the 40 ECU tax and a 40 ECU fine for the non-reported units. Formally, treating $s_{i,t}$ as given, a self-interested risk-neutral seller i in period t chooses the number of reported units $r_{i,t}$ to minimize the expected tax burden

$$V(r_{i,t}, s_{i,t}) = r_{i,t}\tau + \mathbf{1}_{\{r_{i,t}<s_{i,t}\}}\gamma_X(r_{i,t}, s_{i,t})(s_{i,t} - r_{i,t})(f+\tau) \tag{1}$$

where τ is the nominal per-unit tax, f is the fine, $\gamma_X(s_{i,t}, r_{i,t})$ denotes the audit probability in treatment X as a function of $s_{i,t}$ and $r_{i,t}$, and $\mathbf{1}_{\{r_{i,t}<s_{i,t}\}}$ is the indicator taking value one if under-reporting and zero otherwise. The expected effective per-unit tax for seller i in period t is

$$\tau^e(s_{i,t}, r_{i,t}) = V(r_{i,t}, s_{i,t})/s_{i,t}. \tag{2}$$

In the Seller only treatment, $\gamma_{SO}(s_{i,t}, r_{i,t})$ is exogenous. With our parameterization (i.e. $\gamma_{SO}(s_{i,t}, r_{i,t}) = 0.1, \tau = 40$, and $f = 40$), the standard deterrence-model by Allingham and Sandmo (1972) predicts full non-compliance: A risk-neutral and money-maximizing seller optimally reports zero units sold as the marginal benefit from evasion, 40 ECU, exceeds the marginal expected cost $0.1 \cdot (40 + 40) = 8$ ECU.[32]

In SELLER + BUYER and SELLER + BUYERC treatments the expected audit probability depends on the probability with which buyers report trades in addition to $s_{i,t}$ and $r_{i,t}$.[33] To model the probability of audit, we make the simplifying assumption that a buyer's decision to report a given trade is independently and identically distributed across buyers and trades.[34] Hence, each reporting decision is independently Bernoulli distributed, and we denote the buyers' unit reporting probability (i.e. the Bernoulli success probability) by p_{SB} and p_{SBC} in SELLER + BUYER and SELLER + BUYERC, respectively. Given that seller i sells $s_{i,t}$ units in period t, the number of trades buyers report is thus distributed according to $Bin(s_{i,t}, p_X)$, in which $X \in \{\text{SB}, \text{SBC}\}$. With these assumptions, the expected audit probability is given by

$$\gamma_X(s_{i,t}, r_{i,t}) = \gamma(s_{i,t}, r_{i,t}; p_X) = \left[\sum_{k=0}^{r_{i,t}} \binom{s_{i,t}}{k} (p_X)^k (1-p_X)^{s_{i,t}-k}\right] \cdot 0.1 + \left[\sum_{k=r_{i,t}+1}^{s_{i,t}} \binom{s_{i,t}}{k} (p_X)^k (1-p_X)^{s_{i,t}-k}\right] \cdot 0.8. \tag{3}$$

Knowing the quantities $s_{i,t}$, $r_{i,t}$, and p_X permits the calculation of the expected audit probability for seller i according to equation (3).

Seller reporting behavior will thus depend on the expectations about the audit probability which over the periods are predicted to adjust to being correct. By design, the expected audit probability lies between 10% and 80% and is increasing in the extent of the seller's under-reporting and the buyers' reporting rate.[35] Sellers' incentives to report more truthfully and the optimal number of units reported are increasing in the probability of audit, and thus in the reporting rate of the buyers. This implies that reporting a positive number of units sold in the SELLER + BUYER and SELLER + BUYERC treatments may be optimal, depending on how truthfully buyers report trades. In section 5.1 we discuss sellers' best response behavior given the observed buyer reporting rates in treatments SELLER + BUYER and SELLER + BUYERC.

In SELLER + BUYER, buyer reporting is costless and so rational and money-maximizing buyers are indifferent between reporting and not reporting the units they bought. There are therefore two (population) pure strategy equilibria: one where buyers report all their trades and one where buyers report none of their trades. If we assume that buyers report truthfully when (materially) indifferent, e.g. due to a "small" or partial preference for honesty (see Demichelis and Weibull (2008), for instance), incentives for sellers to report truthfully are high as well: 80% audit probability implies an effective unit tax of 64 ECU on unreported units and thus reporting truthfully is optimal for the sellers in this case.[36] If we assume that indifferent buyers are fully non-compliant (i.e. break the tie in favor of the payoff-dominant equilibrium (Harsanyi and Selten, 1988)), selfish buyers would not report their trades, and the expected audit probability would be 10%. In this case, the incentives of risk-neutral and self-interested sellers coincide with those in SELLER ONLY, and reporting zero sold units is optimal.

In SELLER + BUYERC, buyers have to bear a cost if they report a positive number of units, and therefore selfish and money-maximizing buyers are no longer indifferent but rather strictly prefer not reporting at all.[37] Thus, sellers' reporting incentives again coincide with those in SELLER ONLY, and reporting zero sold units is optimal.

Sellers and buyers may have non-pecuniary motives, such as lying aversion and/or image motivation to report truthfully. These are central for the alternative mechanisms discussed in Section 6.[38] In Appendix D, we apply a model based on Gneezy et al. (2018) to understand how lying aversion and image concerns change the predictions for reporting and market outcomes.[39]

C.2. Market prices and quantities

It is well known that double auction market outcomes are generally consistent with the Walrasian equilibrium (see Smith (1962), and Friedman (1993) for a comprehensive survey). Easley and Ledyard (1993) and Gjerstad and Dickhaut (1998) develop theories of

32 This is different from Kleven et al. (2011) where increasing reporting marginally has an effect both on the margin and on the inframarginal under-reporting thus rendering positive reporting optimal even for a self-interested decision maker.

33 Kleven et al. (2011) model the endogeneity of the audit probability in a situation where third-party information is available on a subset of tax items and derive the prediction that reporting will be truthful on third-party reported income items. In that analysis, the existence of third-party information on certain items is taken as given, and the audit probability is common knowledge.

34 None of our predictions rest on this assumption, however. We make it merely to keep the derivation tractable. The crucial feature we need is that the (expected) audit probability is increasing in the seller's under-reporting and in the reporting rate of the buyers. This is a feature that could be achieved with more realistic (and less tractable) assumptions as well. Our central interest is on studying the comparative statics of the seller's incentives between treatments. For this, the comparative statics of the likelihood of audit between treatments play a key role. Correlation in reporting between buyers's units does not preclude the comparative statics.

35 Note that while the ex post *realized* audit probability is either 10% or 80%, ex ante the audit probability is random from the seller's perspective. Thus, ex ante, the expected audit probability is increasing in the seller's under-reporting.

36 It is possible that the effective tax actually *exceeds* 40 ECU if the seller(s) under-report(s) due to overly optimistic expectations regarding buyer reporting, and thereby end up paying both full taxes and fines. This is in fact what we observe in SELLER + BUYER treatment (see Section 5.1).

37 Kotakorpi and Laamanen (2017) argue, in the context of a study analyzing the effect of prefilled income tax returns on compliance, that the fixed cost of filing appears to be a key determinant of the reporting decision.

38 For literature on tax evasion and intrinsic motivation, see e.g. Fortin et al. 2007; Coricelli et al. 2010; Dwenger et al. 2016; Dufwenberg and Nordblom 2022.

39 See also Dufwenberg and Dufwenberg 2018; Abeler et al. 2019; Khalmetski and Sliwka 2019; Barron et al. 2022; Tergiman and Villeval 2023.

price formation in double auction markets with predictions coinciding with the Walrasian predictions. In the model of Gjerstad and Dickhaut (1998) agents' expectations are not merely adaptive but incorporate a better counterfactual understanding than in Easley and Ledyard (1993). We draw from these models and capture the key individual optimality condition which associates ask prices (or accepted bids) and the quantities offered by the seller.

Formally, let $\widetilde{V}(s_{i,t}) \equiv \min_{r_{i,t}} V(r_{i,t}, s_{i,t})$ denote the indirect (dis)utility associated with the reporting decision. The reserve price $\underline{a}^*_{i,X}$ of an experienced seller i is predicted to satisfy

$$
\begin{aligned}
c_i(s^*_{i,X}) + \widetilde{V}(s^*_{i,X}) - \widetilde{V}(s^*_{i,X} - 1) \\
\leq \underline{a}^*_{i,X} \leq a^W_X \leq \\
c_i(s^*_{i,X} + 1) + \widetilde{V}(s^*_{i,X} + 1) - \widetilde{V}(s^*_{i,X}),
\end{aligned} \tag{4}
$$

where $c_i(s)$ is the marginal cost of selling the s:th unit, $s^*_{i,X}$ is the quantity sold by seller i in equilibrium of treatment X, and a^W_X is the Walrasian equilibrium price.[40] We assume that a (experienced) seller correctly anticipates her reporting behavior and the expected effective taxes paid at the reporting stage. The inequalities capture the idea that the seller optimizes the reserve price and sold quantities given the anticipated tax liability and Walrasian market price, so that selling a unit more or less will result in lower profits.

As stated in Section 4, for treatments No tax and Automatic, standard economic theory offers precise quantitative predictions. In the absence of the sales tax, only marginal cost matters for sellers' supply decisions and so the $\widetilde{V}$ terms vanish in equation (4). Supply without tax (solid black line in Fig. 1) equals demand at 17 units in the Walrasian equilibrium, corresponding to a market clearing price between 158 ECU and 162 ECU (interval due to discrete demand and supply). Imposing an automatically collected 40 ECU per-unit sales tax implies a 40 ECU upward shift in the supply curve (dashed black line in Fig. 1). Therefore, market clears with a lower quantity of 13 units and a higher price between 178 ECU and 182 ECU.

In the reporting treatments the supply curve is predicted to shift up by the amount of the effective tax, which depends on the reporting behavior of the traders. In the fully non-compliant equilibrium $\widetilde{V}(s_{i,t}) - \widetilde{V}(s_{i,t} - 1) = 8$ ECU for all $s_{i,t}$. This implies that $\widetilde{V}(s^*_{i,X}) - \widetilde{V}(s^*_{i,X} - 1) = \widetilde{V}(s^*_{i,X} + 1) - \widetilde{V}(s^*_{i,X}) = 8$ ECU in equation (4), and the supply curve shifts up by 8 ECU (blue line in Fig. 1). The fully compliant equilibrium supply (dashed green line) coincides, by construction, with the supply with automatic tax collection.

As discussed in the previous section, effective unit tax in Seller only should provide a lower bound and the unit tax in Automatic an upper bound for the effective unit tax in Seller + Buyer. In the equilibrium where buyers are fully compliant, and hence sellers also are fully compliant, the equilibrium price and quantity are equal to those in Automatic (dashed green line depicting compliant supply in Fig. 1). In the payoff-maximizing collusive equilibrium where buyers are not reporting any units, the equilibrium price and quantity are equal to those in Seller only (solid blue line in Fig. 1). The latter is also the self-interested risk-neutral equilibrium prediction in Seller + BuyerC as self-interested buyers do not report any trades and thus the probabilities of audit coincide in Seller only and Seller + BuyerC.

Appendix D. Theory: lying, social image and incidence

In this section, we present a model of lying behavior with image concerns, which provides an explanation for our non-standard incidence result, as well as our other findings on reporting and market behavior. Our focus here is on explaining the individual variation in sold and reported quantities of the sellers, as well as their reserve prices, in later periods, $t = 11, \dots, 25$, when the sellers are experienced and have well-calibrated expectations and thus stable reserve prices.

In the model, there is individual heterogeneity in reporting behavior, sold quantities, and reserve prices. This heterogeneity is driven by an aversion to lying and by a motivation to appear honest. Abeler et al. (2019) concludes that both of these elements are needed to account for the key patterns in a host of experiments on lying behavior.[41] The novelty of the present model is to apply these models to a market context where lying takes place at the reporting stage, but the image benefits accrue only in the market stage: if one were to supply at an excessively low price, this would be taken as a signal of evasion by other market participants.[42] To our knowledge, these types of models of lying with image concerns have not been previously applied in the tax evasion literature, even though evasion is a prominent real life example of dishonest behavior.

Preferences and timing

Seller i's type is (θ_i, α_i). Types are distributed on $[0, \bar{\theta}] \times [0, \bar{\alpha}]$ according to a cumulative distribution function $F(\cdot,\cdot)$. Parameter θ_i represents a fixed cost of lying, while proneness to image concerns is captured by parameter α_i. The model thus nests the standard type with $(\theta_i, \alpha_i) = (0, 0)$. The type is private information to the seller.

[40] For readability, we denote the cost of k^{th} unit by $c_i(k)$ instead of c_{ik} as in Section 3.

[41] Such models in the context of lying behavior alone have been proposed by Gneezy et al. (2018), Khalmetski and Sliwka (2019). See Attanasi et al. (2019), Bartling et al. (2022) and Tergiman and Villeval (2023) for evidence that intrinsic motivation influences behavior in market and other institutional settings, and Dufwenberg and Dufwenberg (2018) for a model where liars care about the perceptions of others regarding the extent of the lie.

[42] In our setup the units sold are observable to the experimenter and the seller knows this and thus, as in Gneezy et al. (2018) observable game, image plays little role in reporting itself, and only direct lying cost matters.

Following Gneezy et al. (2018), we assume that a seller's preferences depend on three components: monetary earnings, a direct intrinsic cost associated with misreporting (lying), and a social image benefit accrued at the market stage. All components enter the seller's utility additively, and the seller is risk-neutral with respect to monetary earnings. Following Abeler et al. (2019), we assume that there is heterogeneity across intrinsic costs of lying and image concerns. Denoting by s_i and r_i the number of sold and reported units by seller i (we drop the period subindex for simplicity), the intrinsic part of seller i's utility function reads as

$$\alpha_i \rho(a_i) - \mathbf{1}_{\{r_i < s_i\}} \theta_i \tag{5}$$

where $\rho(a_i)$ captures the belief that buyers and other sellers hold about i being an honest type conditional on transaction price a_i to be defined below. $\mathbf{1}_{\{r_i < s_i\}}$ is an indicator function for i's report being a lie. Note that, for simplicity, we assume that the cost of lying does not depend on the extent of under-reporting.[43]

In the tax evasion context, market activities are visible to market participants whereas tax reports are made privately. Thus, the market activities – prices and sold quantities – convey information about the honesty of the seller, not the tax reports directly. As Gneezy et al. (2018), we assume that seller i's image benefit is increasing in the strength of the belief that buyers and other sellers hold about i being an honest type conditional on transaction price a written as

$$\rho(a) = \frac{h(a)}{h(a) + d(a)},$$

where $h(a)$ and $d(a)$ are the probabilities of an honest and a dishonest report when observing prices consistent with reserve price a.

Let us now derive $h(a)$ and $d(a)$. For simplicity, assume that each type plays a pure strategy and chooses a specific supply $s(\theta, \alpha)$, reserve price $\underline{a}(\theta, \alpha)$, and a corresponding reporting strategy $r(s(\theta, \alpha), \theta)$.[44] Let the set of honest types conditional on price a and supplied quantity s be defined as $H(a, s) = \{(\theta, \alpha) : \underline{a}^*(\theta, \alpha) = a, s^*(\theta, \alpha) = s, r^*(s^*(\theta, \alpha), \theta) = s\}$, where $\underline{a}^*, s^*, r^*$ denote optimal strategies. Then the probability of an honest report at price a and quantity s is written as

$$h(a, s) = \iint_{(\theta,\alpha)\in H(a,s)} dF(\theta, \alpha).$$

Likewise, the set of dishonest types conditional on price a and supplied quantity s is defined as $D(a, s) = \{(\theta, \alpha) : \underline{a}^*(\theta, \alpha) = a, s^*(\theta, \alpha) = s, r^*(s^*(\theta, \alpha), \theta) < s\}$, and the probability of a dishonest report at price a and quantity s is written as

$$d(a, s) = \iint_{(\theta,\alpha)\in D(a,s)} dF(\theta, \alpha).$$

Since in our experimental markets prices are salient but the supplied quantities are not, we assume that inferences are made based on observed prices (see Figs. B.1 and B.2 in Appendix B for examples of trading screens). Thus, aggregating over quantities, we have

$$\rho(a) = \frac{h(a)}{h(a) + d(a)},$$

where

$$h(a) = \sum_{s=1}^{4} h(a, s),$$

and

$$d(a) = \sum_{s=1}^{4} d(a, s).$$

Reporting stage

Since none of the other sellers or buyers observe the seller's report, the social image concern does not play any role at the reporting stage.

[43] This is in line with Gneezy et al. (2018), Abeler et al. (2019) and Khalmetski and Sliwka (2019), with the exception that Gneezy et al. (2018) adopt a more complicated formulation and provide evidence that lying cost is weakly increasing and non-convex in the extent of the lie. Yet, their experimental treatments allow them to identify some people whose behavior is consistent with lying cost being increasing and strictly convex in $|s_i - r_i|$ so that the maximal lie is not told. In our case, these predictions would translate into zero sold units being the modal under-report but also sellers inaccurately reporting some sold units. As the evidence in favor of the lying cost having curvature is not particularly strong, we simplify and assume it is constant and thus best captured by a type-specific fixed cost.

[44] In line with Gjerstad and Dickhaut (1998), we assume that, in the later periods where convergence has taken place, all units will be sold at the same price. This is a simplification not entirely consistent with data. We could alternatively use the average price of the units sold by a seller, or the price of the last unit without much complicating the line of argument.

Incorporating lying aversion into the preferences of the sellers in the reporting stage, player i of type (θ_i, α_i) minimizes

$$V(r_i, s_i; \theta_i, \alpha_i) = r_i\tau + \mathbf{1}_{\{r_i<s_i\}}\gamma_X(r_i, s_i)(s_i - r_i)(f + \tau) + \mathbf{1}_{\{r_i<s_i\}}\theta_i \tag{6}$$

where the last term is novel and did not appear in equation (1) which described the preferences of a standard agent only.

In SELLER ONLY the audit probability is constant $\gamma_{SO}(r_i, s_i) = \gamma_{SO} = 0.1$ for all s_i, r_i. In SELLER + BUYER and SELLER + BUYERC treatments, the audit probability is endogenous and depends on whether the seller under-reports relative to her customers. The expected audit probability of seller i selling s_i units and reporting r_i units in treatment $X \in \{\text{SB}, \text{SBC}\}$ is given by equation (3).

We assume that in the market stage, each party correctly anticipates her/his reporting behavior. Thus, we will denote the indirect reporting (dis)utility as $\widetilde{V}(s_i;\theta_i) \equiv \min_{r_i} V(r_i, s_i; \theta_i)$. The expected reporting cost now depends on the lying cost parameter θ_i in addition to expected taxes and fines, and is (weakly) increasing in θ_i.

Market stage

The reserve price $\underline{a}^*_{i,X}$ of an experienced seller i of type (θ_i, α_i) predicted to satisfy

$$\begin{aligned}c_i(s^*_{i,X}) + \alpha_i[\rho(a_X^{s^*_{i,X}-1}) - \rho(\underline{a}^*_{i,X})] + \widetilde{V}(s^*_{i,X}, \theta_i) - \widetilde{V}(s^*_{i,X} - 1, \theta_i)& \\ \le \underline{a}^*_{i,X} \le a^W_X \le & \\ c_i(s^*_{i,X} + 1) + \alpha_i[\rho(\underline{a}^*_{i,X}) - \rho(a_X^{s^*_{i,X}+1})] + \widetilde{V}(s^*_{i,X} + 1, \theta_i) - \widetilde{V}(s^*_{i,X}, \theta_i),&\end{aligned} \tag{7}$$

where $c_i(s)$ is the marginal cost of selling the s:th unit, $s^*_{i,X}$ is the quantity sold by seller i in equilibrium of treatment X, a^W_X is the Walrasian equilibrium price, and a^s_X is the maximum price at which s units could be sold in treatment X. Notice that if lower prices are associated with more tax evasion, then the model predicts that image motivation (higher α_i) is associated with greater reserve price for a given level of supplied quantity, and therefore seller types with higher α_i tend to produce smaller amounts than types with $\alpha_i = 0$. Yet, if θ_i is zero or close to it, and there is no pecuniary motivation to be truthful, the seller type still evades taxes and thus has a lower reserve price and produces more than a truthful type. This is the key mechanism in the predictions that follow.

Treatments

No tax

When there is no tax and no reporting stage, image concerns and direct lying costs play no role, and so the corresponding terms drop out of equation (7). The reserve price of an experienced seller i is predicted to satisfy

$$c_i(s^*_{i,NT}) \le \underline{a}^*_{i,NT} \le a^W_{NT} \le c_i(s^*_{i,NT} + 1).$$

The buyers' induced value schedule and the sellers' induced cost schedule constitute the demand curve and the supply curve, respectively (see Fig. 1), and it is easy to show that the demand and supply price elasticities coincide in equilibrium. In line with the literature following Smith (1962), we find that the average transaction prices and quantities coincide with the theoretical predictions (see Table 2 and Fig. 1).

Automatic tax

In case of automatic taxes, there is no choice but to report all sold units, i.e. $r_i = s_i$ by restriction, and thus $\widetilde{V}(s, \theta_i) - \widetilde{V}(s-1, \theta_i) = \tau$ for all $s > 0$. Consequently, in the market stage, the reserve price of an experienced seller i is predicted to satisfy

$$c_i(s^*_{i,AT}) + \tau \le \underline{a}^*_{i,AT} \le a^W_{AT} \le c_i(s^*_{i,AT} + 1) + \tau.$$

In other words, the predicted equilibrium price lies at the level where the supply curve (aggregated sellers' induced cost schedule shifted up by τ) intersects the demand curve (the buyers' induced value schedule). Since elasticities are unchanged, standard incidence theory predicts that the tax burden is shared equally between buyers and sellers. Indeed, as above in NO TAX, we find that the average transaction prices and quantities coincide with the theoretical predictions, and buyers bear on average 46.8% of the tax burden in periods 11–25 (see Table 2 and Fig. 1).

Seller only

Reporting stage. In SELLER ONLY, heterogeneity in the intrinsic motivation of the sellers enters the picture. The optimal report is given by the solution to the problem

$$\min_{r_i} V(r_i, s_i; \theta_i),$$

or, substituting from equation (6),

$$\min_{r_i}\{r_i\tau + \mathbf{1}_{\{r_i<s_i\}}0.1(s_i - r_i)(f + \tau) + \mathbf{1}_{\{r_i<s_i\}}\theta_i\}.$$

If the seller decides to report truthfully, $r_i = s_i$, and the indirect reporting (dis)utility takes value $s_i\tau$. If the seller under-reports, the fixed lying cost is incurred and the probability of audit is independent of the report at 10%. Conditional on under-reporting, it is optimal not to report any units and the indirect (dis)utility takes value $\theta_i + 0.1s_i(f + \tau)$. This is weakly greater than $s_i\tau$ and hence i reports truthfully[45] if and only if

$$\theta_i \geq s_i(0.9\tau - 0.1f) = 32s_i,$$

which does not hold for $\theta_i = 0$, so that full evasion is optimal for a standard money-maximizing seller. Yet, for θ_i sufficiently high, truth-telling is optimal. Moreover, partial evasion is never optimal (just as in the observable game of Gneezy et al. (2018), see footnote 42). Finally, incentives for tax evasion are increasing in the number of supplied units.

Indeed, the Seller only treatment resembles the die-rolling lying experiment paradigm by Fischbacher and Föllmi-Heusi (2013) in that full lies are optimal for self-interested sellers. A key difference is that in our setup the true state (the number of sold units) is endogenously chosen in the market stage and not exogenously randomized as in the die-rolling task. Furthermore, in our experiment under-reporting comes with an explicit punitive frame and a 10% chance of a monetary fine. Notice also that unlike in a typical Fischbacher and Föllmi-Heusi (2013) setup and in line with the observable game of Gneezy et al. (2018), the true state is public knowledge between the experimenter and the participant while other market participants observe neither the true state nor the report. Thus image motivation, $\alpha_i\rho(\cdot)$, matters equally in the reporting stage of every treatment, having therefore no treatment effect, and only the direct lying cost, θ_i, matters.

Fig. 3 shows that in Seller only about 52% of reports are fully evasive and 18% truthful, consistent with the idea that sellers' intrinsic motivation is heterogeneous.[46] Yet, the remaining 30% of reports indicate partial evasion. The partially evading reports are not consistent with our model but they are consistent with the evidence in the observable game of Gneezy et al. (2018) where 32% of those who lie do so only partially.[47] The model of Gneezy et al. (2018), which our model is based on, also abstracts from the complications of partial lying.

Market stage. As explained, the audit probability is not sufficient to deter tax evasion of a standard seller with $\alpha_i = 0 = \theta_i$. Thus, in the market stage, a standard seller's reserve price and supplied quantity are predicted to satisfy

$$c_i(s^*_{i,SO}) + 0.1(f + \tau) \leq \underline{a}^*_{i,SO} \leq a^W_{SO} \leq c_i(s^*_{i,SO} + 1) + 0.1(f + \tau).$$

Therefore, if all sellers of a session happen to be standard (or their α_i:s and θ_i:s are small), the prediction in Seller only coincides with the prediction in Section C.2. This is what happens in Sessions 14 and 15 in Seller only (see Appendix B).

As is typical for these models, the predicted effects of image motivation are discovered by a fixed point argument: an intrinsically motivated seller with $\theta_i < 32s^*_{i,SO}$ will not report the sold units but cares about social image. Thus, provided that tax-evaders are more likely to agree on lower prices, an image-motivated seller suffers a marginal image cost of $\alpha_i[\rho(\underline{a}^*_{i,SO}(\theta_i, \alpha_i)) - \rho(a^{s^*_{i,SO}+1}_{SO}(\theta_i, \alpha_i))]$ for agreeing on a marginally lower price than her reserve price. Thus, her reserve price for a given unit must be higher than that of the standard type to cover the additional cost, and therefore the predicted number of units will be smaller than for a standard seller. This confirms the idea that types without image motivation are predicted to hold lower reserve prices.

Finally, an intrinsically motivated seller whose fixed lying cost satisfies $\theta_i \geq 32s^*_{i,SO}$ will report all units. She also cares about social image and thus her reserve price must cover the marginal image cost. Moreover, truthful reporting means that the implied tax on the marginal unit, τ, must be covered as well. Thus predicted reserve prices are higher and the predicted number of units sold will be lower than for other seller types.

Consider types who do not have a sufficiently high θ_i to curb tax evasion but who do care about image. They are expected not to charge too low prices to sustain their image in the market place where their actions are observable to other parties. They are rather predicted to mimic the prices of truthful types, thus selling less units than types who evade and do not care about their image. Why is it then that truthful types do not want to deviate upwards and separate to signal their truthfulness? Simply because no buyer is willing to trade at such price and the seller would not be able to sell.

Therefore, if all sellers of a session happen to have high α_i:s and some of them have high θ_i:s, the prediction in such Seller only sessions is that prices are much higher than in the self-interested prediction of Section C.2, even up to the levels of the Automatic treatment. Sellers with a high θ_i report truthfully and those with low θ_i evade fully. This is what happens in Sessions 13 and 16-18 in the Seller only treatment (see Figs. B.18-B.23).

Notice that a Walrasian equilibrium predicts a single price. The law of one price is not literally true in continuous double auctions and the Walrasian equilibrium rather predicts the average price. Some within-session price variation across periods and sellers is observed. It should be noted that in Seller only, a substantial fraction of the price variation is across sessions. On the one hand, there are sessions where prices are high and both full evaders and truthful sellers sell at those prices (sessions 13, 16–18). In these sessions, the fact that evaders do not compete the price down is consistent with the marginal image costs. There are also sessions where competition is tough, prices are bid at No Tax Walrasian equilibrium levels, and evasion is widespread (sessions 14, 15). This

[45] Assuming that indifferent types report truthfully as in Demichelis and Weibull (2008), for instance.

[46] Notice also that the share of truthful reports in Seller only coincides with that in the field experimental settings of Kleven et al. (2011) and Dwenger et al. (2016) suggesting that external validity might not be too severely compromised in the laboratory.

[47] The observable game is the right benchmark at this stage because both the number of units sold and the report are observable to the experimenter. See also Abeler et al. (2019) for evidence in previous experiments.

effect is consistent with random selection of types with little intrinsic motivation who face extreme competitive pressure to bid low and evade. Similar selection effects and variation across sessions is observed in Halliday et al. (2021), in which the effect of selection and competition on lying is explicitly studied. More generally, the Walrasian equilibrium price should be a function of the profile of types of sellers in a given session. Competition is tougher when seller types do not much differ from the standard values. When the α_i:s of the sellers are higher, higher prices can be sustained and sellers with a low θ_i evade taxes. Notice moreover, that the image motivation of the sellers must be supported by correct – albeit counterfactual – beliefs about the types and tax evasion in case prices were competed down to the bottom, exactly as in Halliday et al. (2021).

Although self-interest predicts that the 10% audit probability is not sufficient to deter tax evasion and thus prices will be close to those in the No tax, image-motivation predicts that the average price in Seller only is higher and sold quantities are lower. Moreover, self-interested sellers who are predicted to belong to the fully evading group of sellers, are predicted to (at occasions) trade at a price which undercuts that at which the partially evading or truthful sellers trade. They are also predicted to sell higher quantities.

In line with these predictions, we see that lower transaction prices are more likely in trades by fully evading sellers (Fig. 6): 68% of the sellers trading at an average price lower than or equal to 162 ECU (the No tax equilibrium price) are evading fully, 50% of those trading at a price between 162 and 178 ECU evade fully, and 36% of those selling at a price higher than or equal to 178 ECU (the Automatic equilibrium price) evade fully.

A truthful seller supplies 2.4 units whereas a fully evading seller supplies 3.1 units on average (see also Fig. B.24). These facts are consistent with the predictions that full evaders have lower reserve prices given supplied quantity and they supply more units than truthful sellers.

As the average price of the truthful types is 10 ECU above the (both predicted and observed) No tax price, the truthfully reporting 18% of the sellers are able to pass on average only about a quarter of the 40 ECU unit tax on the buyers. The 52% of the sellers who fully evade pay on average an 8 ECU effective unit tax. As Fig. 6 shows, the fully evading sellers charge somewhat higher prices than the No tax equilibrium price (the left-most vertical dashed line). In fact, for fully evading sellers the average trading price is exactly 8 ECU above the No tax equilibrium, and not 4 ECU which would lead to an equal sharing of the tax burden.

Sellers who under-report fully pass the highest share of the effective tax burden onto buyers.

As this group of seller reports is the largest, in aggregate, the effective tax burden is distorted towards the buyers, unlike in the Automatic. The average tax burden borne by the buyers over periods 11–25 is 62% (see Table 2).

Seller + Buyer

If all agents were of standard type, there would exist a collusive equilibrium where all buyers and sellers report zero units. Yet, a money-maximizing buyer in that equilibrium is indifferent between reporting and not reporting truthfully at the reporting stage. This implies that the equilibrium is not robust to the introduction of buyers who are intrinsically motivated to report truthfully. The direct intrinsic lying costs that we have applied in this section to explain seller reporting can equally well be descriptive of (some of) the buyers. Even a minimal amount of such intrinsic lying costs would render truthful reporting optimal for a buyer.

Consistent with the latter prediction, 81% of the units bought in periods 11–25 are reported by buyers in Seller + Buyer. This yields the audit probability schedule presented in Table 3 in Section 5.1.[48] As we discuss there, when the seller is of the standard type with $\theta_i = 0$ and has rational expectations of buyer reporting, the only cases where under-reporting is optimal is when $s_i = 3$ or $s_i = 4$, with $r_i^* = 2$ and $r_i^* = 3$, respectively. When one or two units are supplied, truthful reporting is optimal even for the standard type. Thus, the behavioral model, with the implied intrinsic motivation for truthfulness, increases the explanatory power of the model on the seller side by capturing the higher than predicted truthful reporting conditional on $s_i = 3$ or $s_i = 4$, and by explaining the high reporting rate on the buyer side.

Since seller-reporting is predicted to be at a high level, so are paid taxes. Even in the cases where the standard type optimally under-reports, the incentive to do so is quite modest: the expected sum of taxes and fines is quite close to the tax paid by truthful reporters. This may cause even the standard type seller, for whom under-reporting is optimal, to sub-optimally report truthfully.[49] Thus, the fraction of sellers who under-report is predicted to be small, and the reserve prices of standard and intrinsically motivated types are predicted to largely overlap. Therefore, no major image effects of charging lower or higher prices are predicted unlike in Seller only.[50] Thus, the predicted supply curve in Seller + Buyer approximates or at most slightly undercuts that in Automatic. Moreover, due to equally elastic equilibrium supply and demand, the effective tax burden is expected to be shared in equal proportions, just as predicted by the standard tax incidence model.

These predictions coincide with what is observed in the data (see Figs. 2, 3, and 4). About 78% of sellers' reports are truthful and only 6% of reports under-report by more than one unit. Average prices of evaders are not significantly different from those of

[48] This approach of course misses the session-specific variation in buyer reporting behavior. Although sellers adapt their reporting and sales behavior based on buyers' behavior in their own session (recall that we focus on periods 11 to 25 where sellers have had ample opportunity to learn from the buyers' behavior in the session), we see the model as tractable and simple, and as one that helps the reader in understanding the central trade-offs.

[49] When the incentive profile over different reports is relatively flat, mistakes in the form of slightly sub-optimal reports become more likely (as in quantal response models (McKelvey and Palfrey, 1995, 1998)).

[50] In other words, Equation (7) looks very similar for Seller+Buyer treatment as for Automatic treatment. This happens because the difference between the ρ-terms is minimal for Seller+Buyer (and zero for Automatic). Moreover, $\widetilde{V}(s^*_{i,SB},\theta_i) - \widetilde{V}(s^*_{i,SB}-1,\theta_i) \approx \tau$ for many i and $\widetilde{V}(s^*_{i,SB},\theta_i) - \widetilde{V}(s^*_{i,SB}-1,\theta_i) = \tau$ for some i, whereas $\widetilde{V}(s^*_{i,AT},\theta_i) - \widetilde{V}(s^*_{i,AT}-1,\theta_i) = \tau$ for all i.

truthful reporters (also because under-reporting is so rare, see Figs. 6 and B.16). Buyers bear on average 46% of the tax burden in periods 11–25, not significantly different from the predicted 50% (see Table 2 and Fig. 5).

Seller + BuyerC

As reporting is costly for buyers, a standard money-maximizing buyer prefers not to report any units. Nevertheless, as the cost of reporting is small, some intrinsic lying aversion is sufficient to make truthful reporting optimal.

Consistent with the existence of intrinsically motivated buyers, the reporting rate of buyers is 43% which is both significantly higher than zero and significantly lower than in Seller + Buyer.[51] In Section 5.1, we noted that with $p_{SBC} = 0.43$ the implied audit probabilities are such that for the standard seller type with $\theta_i = 0$ and rational expectations regarding buyer reporting, it is optimal to under-report by one unit if the produced amount is three or less than three, and under-report by two units if the amount supplied is four. Further under-reporting is deterred by the high implied marginal effect of expected effective tax (i.e. expected sum of taxes and fines) which is due to the negative effects on infra-marginal units. The expected sum of the fine and the tax for the first under-reported unit is almost as high as the unit tax. Thus, the reserve prices of the standard type are slightly below but close to the reserve prices in Automatic. As extrinsic incentives already fairly effectively deter evasion, some intrinsic motivation is sufficient to deter any under-reporting, and thus some of the intrinsically motivated types are predicted to report truthfully in Seller + BuyerC.

Turning to our evidence, in Seller + BuyerC, in which a standard type should under-report and some intrinsically motivated types should report truthfully, about 42% of the sellers' reports are truthful. The behavioral model incorporating lying aversion and image motivation thus raises the explanatory power of the model by capturing the high rates of truthful reporting on both the seller and the buyer side.

The observed under-reporting is more common in Seller + BuyerC than in Seller + Buyer. Furthermore, under-reporting by one unit is the most common form of under-reporting unlike in Seller only (see Fig. 3).[52] These patterns are in line with both the self-interested best-response to buyer behavior and the behavioral model with lying aversion and image motivation. About 30% of the sellers under-report by two or more units. The share of sub-optimal reporting decisions is higher in Seller + BuyerC than in Seller + Buyer (see Fig. B.15).[53]

Reserve prices of all reporting profiles are predicted to be higher than in Seller only. The reserve prices of under-reporting sellers are predicted to be slightly lower and/or the supplied quantities slightly higher than those of truthful sellers. Thus, some tendency for lower prices to signal under-reporting is predicted but this effect should be small and thus have little effect through image motivation on reserve prices. In fact, in our data, the fraction of truthful reports is constant between 41% and 44% across price brackets below 162 (the No Tax equilibrium price), 162–170, 170–178, and above 178 (the Automatic equilibrium price), and the distribution of average prices of full and partial evaders are not significantly different from those of truthful reporters (Fig. 6).

As no major under-reporting is optimal and supply curves approximate those in the Automatic treatment (with equally elastic supply and demand), the tax burden is expected to be shared in equal proportions, just as predicted by the standard tax incidence model. In Seller + BuyerC, buyers bear 37% of the effective tax burden (see Table 2 and Fig. 5), which is not statistically significantly different from 50%, the prediction of the standard theory. Therefore, sellers have no benefit of tax evasion opportunities but rather bear a greater effective tax burden. To some extent, this is due to the sub-optimal evasion behavior, which results in high and sub-optimal expected effective tax rates (see footnote 53).

Appendix E. Experimental instructions

Translated instructions for the Seller + BuyerC *treatment*

General instructions

This session is part of an experiment on decision making. If you follow the instructions and make good decisions, you can earn money. Your earnings will be paid to you privately in cash. How much you earn depends on your own decisions and the decisions of other participants.

There are 10 persons in this room taking part in this experimental session. You are not allowed to communicate with others during the experiment. We kindly ask that you read these instructions very carefully. If you have questions, please raise your hand and the experimenter will come to you and answer your questions. Your decisions throughout the experiment, and your earnings will be treated confidentially.

You can earn money in this experiment. Your earnings depend on your and other participants' decisions and possibly chance. During the experiment, your earnings are calculated in an experimental currency called ECU ("Experimental Currency Unit"). At the end of the experiment your earnings will be converted to EURO so that 500 ECU = 1 EUR, and paid to you in private along with a 5 EUR participation fee.

The experiment

[51] It is also at about the same level as non-zero reporting by sellers in Seller only where extrinsic incentives to report are also small.

[52] See also Fig. B.17 which displays trading prices by sold units and reporting decision in Seller + BuyerC treatment.

[53] The sub-optimal behavior can be partially due to the highly stochastic feedback which makes it hard for sellers to estimate the expected fine for each level of under-reporting. The number of a seller's trades reported by her customers is Binomially distributed with equal success and failure rates. Recall that only individual audit outcomes are observable to sellers after each round, not the reports nor the reporting rate of the buyers.

Roles

At the beginning of the experiment, the computer will randomly assign 5 participants to the role of "seller" and 5 participants to the role of "buyer". Therefore, you will either be a buyer or a seller. Your role as buyer or seller will remain the same throughout the experiment. You will only know your own role, and not the roles of others.

Overview

The experiment consists of 3 practice periods and 25 actual periods. Only the 25 actual periods affect your earnings. At the beginning of a decision period there is a market phase, during which sellers and buyers trade a fictitious good in a market place. As a buyer, you can buy units of the fictitious good, and as a seller, you can sell units.

You can earn ECU by trading in the market place, and your earnings depend on your, and other participants' decisions. Sellers are liable to pay a 40 ECU unit tax on each unit they sell. The tax is the same for all sellers and is due after each market stage. All "tax revenue" collected in the experiment is donated to the Finnish State.

The market

Basics

The market place opens for trading for 100 seconds at the beginning of each period. In the market traders trade a fictitious good. Each seller can sell up to 4 units, and each buyer can buy up to 4 units of the fictitious good. Trade is conducted through a trading screen.

Goods, costs and values

If you are a seller, at the beginning of the experiment, you will be randomly assigned the production costs ("costs" from now on) for 4 units of the fictitious good. These units are denoted "Good 1", "Good 2", "Good 3" and "Good 4". The cost of Good 1 is lower than the cost of Good 2, the cost of Good 2 is lower than the cost of Good 3 and the cost of Good 3 is lower than the cost of Good 4. These costs will remain the same to you throughout the experiment. The costs of each seller differ from the costs of other sellers' goods. Each seller only knows her own costs.

If you are a buyer, at the beginning of the experiment, you will be randomly assigned the values for 4 units of a fictitious good. These goods are denoted "Good 1", "Good 2", "Good 3" and "Good 4". The value of Good 1 is higher than the value of Good 2, the value of Good 2 is higher than the value of Good 3 and the value of Good 3 is higher than the value of Good 4. These values will remain the same to you throughout the experiment. The values of each buyer differ from the values of other buyers' goods. Each buyer only knows her own values.

Asks, bids, and trading

Sellers can make offers to sell and buyers can make offers to buy during the market phase. The lowest standing offer to sell and the highest standing offer to buy are visible to everyone on their trading screen. The screen also states whether you are a seller or a buyer, how much time is left in the market phase and the costs or values that you were assigned for each of your 4 goods.

Each seller first has to sell Good 1 (the good with the lowest cost), then Good 2, then Good 3 and finally Good 4. Accordingly, each buyer first has to buy Good 1 (the good with the highest value), then Good 2, then Good 3 and finally Good 4.

Sellers cannot sell goods at a price that is lower than the cost for the respective good. Buyers cannot buy units at a price that exceeds the value for the respective good.

Sellers can post offers to sell any time during the market phase but each offer to sell has to be lower than the current lowest offer to sell on the market. Accordingly, buyers can post offers to buy any time during the market phase but each offer to buy has to be higher than the current highest bid on the market.

A transaction takes place when either a seller accepts an offer to buy or a buyer accepts an offer to sell. The transaction price for the good then equals the accepted offer to sell or buy.

See Image 1: Example of a seller's trading screen, and Image 2: Example of a buyer's trading screen.

Here screenshot: Image 1. Example of a seller's trading screen

The upper bar of the trading screen displays the current period and how much time is left for trading. Seller's costs, gross profits, number of goods sold in the current period and a reminder of the per-unit tax are displayed on the left in the middle section. Note that the costs and tax in this example are not the same as those in this experiment. The trading prices are displayed on the right in the order in which the goods have been traded. The lower part of the screen shows the current standing offer to sell and current standing offer to buy. The seller can accept an offer to buy by pressing the "Sell at this price" button. To post a lower offer to sell, the seller has to write the offer in the empty field next to the "Make a lower offer" button and press the button.

Here screenshot: Image 2. Example of a buyer's trading screen

The upper bar of the trading screen displays the current period and how much time is left for trading. Buyer's value, gross profits and the number of goods bought in the current period are displayed on the left in the middle section. Note that the values in this example are not the same as those in this experiment. The trading prices are displayed on the right in the order in which the goods have been traded. The lower part of the screen shows the current standing offer to sell and current standing offer to buy. The buyer can accept an offer to sell by pressing the "Buy at this price" button. To post a higher offer to buy, the buyer has to write the offer in the empty field next to the "Make a higher offer" button and press the button.

Gross earnings from trading

Goods that are not bought or sold do not yield profits or losses. Gross profit from each traded good is the following:

Sellers

Gross profit from selling Good = Trading price of Good 1 - Cost of Good 1
Gross profit from selling Good 2 = Trading price of Good 2 - Cost of Good 2
Gross profit from selling Good 3 = Trading price of Good 3 - Cost of Good 3
Gross profit from selling Good 4 = Trading price of Good 4 - Cost of Good 4

Buyers

Gross profit from buying Good 1 = Value of Good 1 - Trading price of Good 1
Gross profit from buying Good 2 = Value of Good 2 - Trading price of Good 2
Gross profit from buying Good 3 = Value of Good 3 -Trading price of Good 3
Gross profit from buying Good 4 = Value of Good 4 - Trading price of Good 4

Gross earnings from trading equal the sum of gross profits.

Reporting of trades

After the trading phase each seller and buyer makes a decision concerning the reporting of the goods he traded in the current period.

Seller's reporting decision

Sellers are liable to pay a per-unit tax (40 ECU) for each good they trade, and the sum of taxes payable is determined by the number of trades a seller *reports* unless the report is checked for accuracy (see "The effect of reports" below). A seller can report any number between zero and the number of goods he traded in the current period. The reporting decision is sent by pressing the "OK" button.

Buyer's reporting decision

A buyer makes a reporting decision concerning the goods he bought in the current period. A buyer pays a 10 ECU reporting cost if he reports one or more goods he bought in the current period. A buyer reports by ticking a box next to every good he bought. The reporting decision is sent by pressing the "OK" button.

The effect of reports

Whether a seller's reported number of trades equals the number of goods she actually sold in the current period can be checked. The probability of a seller's report being checked is determined as follows:

- In the basic case the seller's report is checked for accuracy with a probability of **10%**.
- In addition, the seller's and her trading partners' (buyers who bought from her) reports are cross-checked. If there is a mismatch between the reports so that the number of goods the seller reported as sold is lower than the number of goods bought from the seller reported by her trading partners, the probability that the seller's report is checked for accuracy is **80%**. If, instead, the number of goods reported by the seller is larger than the number reported by her trading partners, the probability of the check for accuracy is not affected.

The probability of a seller's report being checked is not affected by the seller's possible previous checks nor whether other sellers' reports are checked in the current period.

Example: Seller A sold all her 4 goods, but reports selling 1 good.

a If at most one of Seller A's trading partners reports having bought a good sold by her, the probability that Seller A's report is checked for accuracy is 10% (one in ten).
b If two or more of Seller A's trading partners report having bought goods sold by her, the probability that Seller A's report is checked for accuracy is 80% (eight in ten).

Calculation of net earnings

Sellers' net earnings

After the reporting phase the screen displays how many goods you sold and your gross profits. Your **net earnings** depend on the taxes you pay and possible fines. After the reporting phase, one of the following takes place:

1 **The seller's report is not checked for accuracy:** In this case the seller's profit after taxes, i.e. net earnings, equals the sum of gross profits earned in the current period minus taxes. Taxes payable equal the number of goods reported by the seller times the 40 ECU tax:
*Net earnings = sum of gross profits - (reported number of goods sold * 40 ECU tax)*

2 **The seller's report is checked for accuracy:** In this case the seller's profit after taxes, i.e. net earnings, equals the sum of gross profits earned in the current period minus taxes and possible fines. Taxes payable equal the number of goods actually sold by the seller times the 40 ECU tax. If the number of goods reported by the seller is smaller than the number of goods he actually sold, the seller has to pay a fine that equals the per unit tax (40 ECU) for each good he did not report in addition to the missing taxes:
*Net earnings = sum of gross profits - (actual number of goods sold * 40 ECU tax) - (number of goods not reported * 40 ECU tax)*

Buyers' net earnings

After the reporting phase the screen displays how many goods you bought and your earnings is one of the following. Buyer's net earnings depend on the reporting cost. Buyer's net earnings is either of the following:

1 **Buyer's net earnings in case she does not report any trades:**
Net earnings = sum of gross profits
2 **Buyer's net earnings in case she reports at least one of his trades:**
Net earnings = sum of gross profits - 10 ECU reporting cost

Example 1: Seller's earnings

Seller A sold 2 goods. The cost of Good 1 is 112 ECU and the trading price 200 ECU, and the cost of Good 2 is 140 ECU and the trading price 171 ECU. The net earnings of Seller A:

i If Seller A reports both trades: 200 - 112 + 171 - 140 - 2*40 = **39 ECU**
ii If Seller A reports 0 trades and the report is not checked for accuracy: 200 - 112 + 171 - 140 = **119 ECU**
iii If Seller A reports 0 trades and the report is checked for accuracy: 200 - 112 + 171 - 140 - 2*40 - 2*40 = **-41 ECU**

Example 2: Buyer's earnings

Buyer B buys 3 goods. The value of Good 1 is 213 ECU and trading price 180 ECU, the value of Good 2 is 118 and trading price 100, and the value of Good 3 is 110 and trading price 105 ECU. Buyer B's net earnings:

i If Buyer B reports 0 trades: 213 - 180 + 118 - 100 + 110 - 105 = **56 ECU**
ii If Buyer B reports one or more trades: 213 - 180 + 118 - 100 + 110 - 105 - 10 = **46 ECU**

Payoffs

The first 3 periods are practice periods during which you cannot earn money. The 25 periods after the practice periods are payoff relevant, and your total earnings from the experiment consist of your net earnings from these periods and a 5 EUR participation fee.

If the sum of your net earnings is negative, you will be paid the participation fee, so you cannot make losses in this experiment and you will earn at least 5 EUR. Your total earnings will be paid to you in cash after the experiment.

Final note

The experiment ends after 28 periods. After this, we kindly ask you to fill out a short questionnaire while we prepare the payments. All information gathered in the questionnaire, as well as other data gathered in the experiment, will be handled confidentially and used solely for scientific research. After you have completed the questionnaire we ask that you stay seated until we invite you to collect your payment.

Appendix F. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jebo.2024.01.021.

References

Abeler, J., Nosenzo, D., Raymond, C., 2019. Preferences for truth-telling. Econometrica 87 (4), 1115–1153.
Abraham, M., Lorek, K., Richter, F., Wrede, M., 2017. Collusive tax evasion and social norms. Int. Tax Public Finance 24 (2), 179–197.
Agranov, M., Buyalskaya, A., 2022. Deterrence effects of enforcement schemes: an experimental study. Manag. Sci. 68 (5), 3573–3589.
Allingham, M.G., Sandmo, A., 1972. Income tax evasion: a theoretical analysis. J. Public Econ. 1 (3–4), 323–338.
Alm, J., 2019. What motivates tax compliance? J. Econ. Surv. 33 (2), 353–388.
Alm, J., Deskins, J., McKee, M., 2009. Do individuals comply on income not reported by their employer? Public Finance Rev. 37 (2), 120–141.
Alm, J., Jackson, B., McKee, M., 1992. Institutional uncertainty and taxpayer compliance. Am. Econ. Rev. 82 (4), 1018–1026.
Alm, J., Malézieux, A., 2021. 40 years of tax evasion games: a meta-analysis. Exp. Econ. 24 (3), 699–750.
Alm, J., Sennoga, E.B., 2010. Mobility, competition, and the distributional effects of tax evasion. Natl. Tax J. 63 (4, Part 2), 1055–1084.
Andreoni, J., 2006. Leadership giving in charitable fund-raising. J. Public Econ. Theory 8 (1), 1–22.
Antinyan, A., Asatryan, Z., 2019. Nudging for tax compliance: a meta-analysis. ZEW Discussion Paper 19–055.
Attanasi, G., Rimbaud, C., Villeval, M.C., 2019. Embezzlement and guilt aversion. J. Econ. Behav. Organ. 167, 409–429.
Balafoutas, L., Beck, A., Kerschbamer, R., Sutter, M., 2015. The hidden costs of tax evasion: collaborative tax evasion in markets for expert services. J. Public Econ. 129, 14–25.
Barron, K., Kajackaite, A., Saccardo, S., 2022. Image concerns and lying behavior. Available at SSRN. https://ssrn.com/abstract=4204681.
Barron, K., Nurminen, T., 2020. Nudging cooperation in public goods provision. J. Behav. Exp. Econ. 88, 101542.

Bartling, B., Fehr, E., Özdemir, Y., 2022. Does market interaction erode moral values? Rev. Econ. Stat., 1–10.
Bayer, R., Cowell, F., 2009. Tax compliance and firms' strategic interdependence. J. Public Econ. 93, 1131–1143.
Benistant, J., Galeotti, F., Villeval, M.C., 2022. Competition, information, and the erosion of morals. J. Econ. Behav. Organ. 204, 148–163.
Benzarti, Y., 2020. How taxing is tax filing? Using revealed preferences to estimate compliance costs. Am. Econ. J.: Econ. Policy 12 (4), 38–57.
Bjørneby, M., Alstadsæter, A., Telle, K., 2021. Limits to third-party reporting: evidence from a randomized field experiment in Norway. J. Public Econ. 203, 104512.
Bø, E.E., Slemrod, J., Thoresen, T.O., 2015. Taxes on the Internet: deterrence effects of public disclosure. Am. Econ. J.: Econ. Policy 7 (1), 36–62.
Borck, R., Engelmann, D., Müller, W., Normann, H.-T., 2002. Tax liability-side equivalence in experimental posted-offer markets. South. Econ. J., 672–682.
Bott, K.M., Cappelen, A.W., Sørensen, E.Ø., Tungodden, B., 2020. You've got mail: a randomized field experiment on tax evasion. Manag. Sci. 66 (7), 2801–2819.
Carrillo, P., Pomeranz, D., Singhal, M., 2017. Dodging the taxman: firm misreporting and limits to tax enforcement. Am. Econ. J. Appl. Econ. 9 (2), 144–164.
Casal, S., Mittone, L., 2016. Social esteem versus social stigma: the role of anonymity in an income reporting game. J. Econ. Behav. Organ. 124, 55–66.
Coricelli, G., Joffily, M., Montmarquette, C., Villeval, M.C., 2010. Cheating, emotions, and rationality: an experiment on tax evasion. Exp. Econ. 13 (2), 226–247.
Cox, J.C., Rider, M., Sen, A., 2018. Tax incidence: do institutions matter? An experimental study. Public Finance Rev. 46 (6), 899–925.
Demichelis, S., Weibull, J.W., 2008. Language, meaning, and games: a model of communication, coordination, and evolution. Am. Econ. Rev. 98 (4), 1292–1311.
Doerr, A., Necker, S., 2021. Collaborative tax evasion in the provision of services to consumers: a field experiment. Am. Econ. J.: Econ. Policy 13 (4), 185–216.
Doerrenberg, P., Duncan, D., 2019. How does firm tax evasion affect prices? https://www.bwl.uni-mannheim.de/media/Lehrstuehle/bwl/Doerrenberg/How_does_firm_tax_evasion_affect_prices.pdf.
Dufwenberg, M., Dufwenberg, M.A., 2018. Lies in disguise–a theoretical analysis of cheating. J. Econ. Theory 175, 248–264.
Dufwenberg, M., Nordblom, K., 2022. Tax evasion with a conscience. J. Public Econ. Theory 24 (1), 5–29.
Dwenger, N., Kleven, H., Rasul, I., Rincke, J., 2016. Extrinsic and intrinsic motivations for tax compliance: evidence from a field experiment in Germany. Am. Econ. J.: Econ. Policy 8 (3), 203–232.
Dwenger, N., Treber, L., 2022. Shaming for tax enforcement. Manag. Sci. 68 (11), 8202–8233.
Easley, D., Ledyard, J.O., 1993. Theories of price formation and exchange in double oral auctions. In: Friedman, D., Rust, J. (Eds.), The Double Auction Market: Institutions, Theories and Evidence. Addison-Wesley.
Eerola, E., Kosonen, T., Kotakorpi, K., Lyytikäinen, T., 2019. Tax compliance in the rental housing market: Evidence from a field experiment. Technical report, VATT Institute for Economic Research, Working Paper No. 122.
Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10 (2), 171–178.
Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in disguise - an experimental study on cheating. J. Eur. Econ. Assoc. 11 (3), 525–547.
Fortin, B., Lacroix, G., Villeval, M.-C., 2007. Tax evasion and social interactions. J. Public Econ. 91 (11–12), 2089–2112.
Friedman, D., 1993. The double auction market institution: a survey. In: Friedman, D., Rust, J. (Eds.), The Double Auction Market: Institutions, Theories and Evidence. Addison-Wesley.
Gjerstad, S., Dickhaut, J., 1998. Price formation in double auctions. Games Econ. Behav. 22 (1), 1–29.
Gneezy, U., Kajackaite, A., Sobel, J., 2018. Lying aversion and the size of the Lie. Am. Econ. Rev. 108 (2), 419–453.
Greiner, B., 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1 (1), 114–125.
Grossman, Z., 2015. Self-signaling and social-signaling in giving. J. Econ. Behav. Organ. 117, 26–39.
Halliday, S.D., Lafky, J., Wilson, A., 2021. Honesty at the margin: Experiments on the dynamics of deception.
Harju, J., Kosonen, T., Slemrod, J., 2020. Missing miles: evasion responses to car taxes. J. Public Econ. 181, 104108.
Harju, J. Kotakorpi, K. Matikka, T. Nivala, A., 2023. Risk-based tax audits and firm performance, Technical report. Mimeo.
Harsanyi, J.C., Selten, R., 1988. A General Theory of Equilibrium Selection in Games. MIT Press.
Imas, A., 2014. Working for the "warm glow": on the benefits and limits of prosocial incentives. J. Public Econ. 114, 14–18.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica, 263–291.
Khalmetski, K., Sliwka, D., 2019. Disguising lies—image concerns and partial lying in cheating games. Am. Econ. J. Microecon. 11 (4), 79–110.
Kleven, H.J., 2014. How can Scandinavians tax so much? J. Econ. Perspect. 28 (4), 77–98.
Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S., Saez, E., 2011. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. Econometrica 79 (3), 651–692.
Kopczuk, W., Marion, J., Muehlegger, E., Slemrod, J., 2016. Does tax-collection invariance hold? Evasion and the pass-through of state diesel taxes. Am. Econ. J.: Econ. Policy 8 (2), 251–286.
Kotakorpi, K., Laamanen, J.-P., 2017. Prefilled income tax returns and tax compliance: Evidence from a natural experiment. VATT and Tampere University. Mimeo.
Kotakorpi, K., Nurminen, T., Miettinen, T., Metsälampi, S., 2018. Pre-analysis plan: tax evasion in the lab - a comparison of reporting institutions in a competitive market. https://osf.io/bkeca/?view_only=8e1c00932b6d402db820615333852be5.
Luttmer, E.F.P., Singhal, M., 2014. Tax morale. J. Econ. Perspect. 28 (4), 149–168.
Malézieux, A., 2018. A practical guide to setting up your tax evasion game. J. Tax Adm. 4 (1), 107–127.
McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10 (1), 6–38.
McKelvey, R.D., Palfrey, T.R., 1998. Quantal response equilibria for extensive form games. Exp. Econ. 1 (1), 9–41.
Naritomi, J., 2019. Consumers as tax auditors. Am. Econ. Rev. 109 (9), 3031–3072.
Okat, D., 2016. Deterring fraud by looking away. Rand J. Econ. 47 (3), 734–747.
Paulus, A., 2015. Income underreporting based on income expenditure gaps: Survey vs tax records. ISER Working Paper 2015-15.
Perez-Truglia, R., Troiano, U., 2018. Shaming tax delinquents. J. Public Econ. 167, 120–137.
Pomeranz, D., 2015. No taxation without information: deterrence and self-enforcement in the value added tax. Am. Econ. Rev. 105 (8), 2539–2569.
Pomeranz, D., Vila-Belda, J., 2019. Taking state-capacity research to the field: insights from collaborations with tax authorities. Annu. Rev. Econ. 11, 755–781.
Ruffle, B.J., 2005. Tax and subsidy incidence equivalence theories: experimental evidence from competitive markets. J. Public Econ. 89 (8), 1519–1542.
Slemrod, J., 2007. Cheating ourselves: the economics of tax evasion. J. Econ. Perspect. 21 (1), 25–48.
Slemrod, J., Collins, B., Hoopes, J.L., Reck, D., Sebastiani, M., 2017. Does credit-card information reporting improve small-business tax compliance? J. Public Econ. 149, 1–19.
Slemrod, J., Gillitzer, C., 2014. Tax Systems. MIT Press.
Slemrod, J., Rehman, O.U., Waseem, M., 2022. How do taxpayers respond to public disclosure and social recognition programs? Evidence from Pakistan. Rev. Econ. Stat. 104 (1), 116–132.
Smith, V.L., 1962. An experimental study of competitive market behavior. J. Polit. Econ. 70 (2), 111–137.
Tergiman, C., Villeval, M.C., 2023. The way people Lie in markets: detectable vs. deniable lies. Manag. Sci. 69 (6), 3340–3357.
Tonin, M., Vlassopoulos, M., 2014. An experimental investigation of intrinsic motivations for giving. Theory Decis. 76 (1), 47–67.