Estimating the Under-Reporting of Cigarette Production in Pakistan

Tobacconomics Working Paper Series

Muhammad Sabir, Wasim Saleem, and Muhammad Asif Iqbal

Paper No. 22/9/2

September 2022

Correspondence to: Muhammad Sabir,

Social Policy and Development Centre (SPDC), Pakistan

muhammadsabir@hotmail.com

Suggested citation: Sabir, M., Saleem, W., & Iqbal, M.M. (2022). Estimating the Under-reporting of Cigarette Production in Pakistan (Tobacconomics Working Paper No.22/9/2). SPDC. https://tobacconomics.org/research/estimating-the-under-reporting-of-cigarette-production-in-pakistan-working-paper-series

Acknowledgements: This research is funded by the University of Illinois Chicago's (UIC) Institute for Health Research and Policy to conduct economic research on tobacco taxation in Pakistan. UIC is a partner of the Bloomberg Philanthropies' Initiative to Reduce Tobacco Use. The views expressed in this document cannot be attributed to, nor can they be considered to represent, the views of UIC, the Institute for Health Research and Policy, or Bloomberg Philanthropies.

Abstract

Background

Tobacco taxation is a widely-used policy tool to reduce tobacco consumption globally. In Pakistan, however, policy makers have fallen short of tapping the full potential of cigarette taxation, and inefficiencies in the tax collection system persist. The national tax collection agency, the Federal Board of Revenue (FBR), relies on manufacturers' self-declaration of production to determine their tax liability. In the absence of an effective audit system, this mechanism contains a built-in opportunity for under-reporting production to evade taxes and thus causes inefficiencies in tax collection. In this context, the study's goal is to quantify the under-reporting of cigarette production and assess the extent of tax evasion by the cigarette manufacturing industry.

Methodology

A three-step methodology is adopted to calculate the magnitude of under-reported cigarette production. The first step is to construct a database from the financial statements of three cigarette manufacturing companies listed on the Pakistan Stock Exchange and from other relevant secondary sources. The second step is to estimate a theoretically consistent production function using the database. The third and final step is to compute the aggregate annual values of production, convert them into the number of cigarette sticks, and compute the magnitude of under-reported cigarettes by taking the difference between reported and estimated production.

Results

Estimates suggest that the three listed firms under-reported the production of more than 12.8 billion cigarette sticks from 2018-19 to 2020-21, translating to an average of 8.1 percent under-reporting, which resulted in estimated tax evasion of PKR29.5 billion during the three years.

Conclusions

Pakistan's large fiscal imbalances necessitate increased tax revenues, and tobacco taxation provides one avenue for this. The analysis provides evidence of a significant level of under-reporting of cigarette production by the firms, which has caused a substantive revenue loss to the national exchequer. Also, the under-reported production permits the industry more flexibility in pricing so that they can charge less if they choose to attract or retain customers.

JEL Codes: L66, H23, and H26

Keywords: Tobacco, Tobacco Taxes, Tax Under-reporting (evasion)

Table of Contents

Introduction

Global evidence has established that tobacco taxation is an effective policy tool for reducing tobacco consumption (Chaloupka et al., 2019; Chaloupka et al., 2012). However, tobacco taxation policies in Pakistan have fallen short of this goal – with a high burden of tobacco use where nearly one in every five adults uses tobacco in some form.[1] The average excise tax share in Pakistan is 40.9 percent of the retail price of cigarettes,[2] much lower than the widely-accepted benchmark of 70 percent. The cause of this situation can be attributed to lax tobacco taxation policy and inefficiencies in the tax collection system. Moreover, a common argument by the cigarette industry is that higher taxes encourage illicit trade. This line of argument is sometimes even endorsed by the national tax collection agency, the Federal Board of Revenue (FBR).[3]

In Pakistan, domestically produced cigarettes are subject to two major indirect taxes – the Federal Excise Duty (FED) and the General Sales Tax (GST). The FED accounts for almost 80% of the revenue from the sector. Both taxes are collected at the production stage. The cigarette manufacturing sector in Pakistan is an example of an imperfect market, with only three companies dominating domestic production. Despite the small number of large manufacturers, the Federal Board of Revenue (FBR) relies on manufacturers' self-declaration of production to determine their tax liability. In the absence of an effective audit/monitoring system, this mechanism contains a built-in opportunity for under-reporting production to evade taxes and thus causes inefficiencies in tax collection. For example, Ross (2018) cited Brazil and Vietnam as countries where the implementation of a track and trace system, as well as legislation and improved coordination, aided in reducing the under-reporting of domestic production. In Pakistan, the self-reporting mechanism tends to encourage under-reporting and manipulation of the production numbers, which is then used to influence tax policy in favour of the industry. For instance, Iqbal et al. (2020) estimated yearly variations in the magnitude of under-reporting from 2015-16 to 2017-18 using supply and production functions. The results show a sharp increase of more than 40 percent in the under-reporting of cigarettes during 2016-17. This enormous magnitude of under-reporting was successfully used to convince tax authorities to introduce a third tier in the FED structure with a lower tax rate.[4]

[1] Authors' estimates based on GATS 2014 and population projection for 2021.
[2] Budget 2022-23: Revenue and health implications of excise tax increase on cigarettes, Policy Brief, Social Policy and Development Centre (SPDC), July 2021.
[3] For instance, this argument was made by FBR before a special committee of the Senate of Pakistan (Source: Report of the Special Committee on Causes of Decline in Tax Collection of Tobacco Sector, 2018)
[4] In 2013, a two-tier structure of specific FED based on range of retail prices (exclusive of GST) was adopted. Until 2016-17, the two-tier system remained intact with annual upward revision of tax rates of both the tiers. However, in 2017, the federal government introduced a three-tier excise duty structure for cigarettes – with a new tier for the low-priced brands. The tax rate applicable on the new tier was reduced by 48 percent.

Linking the production of cigarettes to its inputs, particularly raw materials, is important to the dynamics of cigarette production and to isolate it from reporting anomalies. An initial attempt was made by Iqbal et al. (2020) to estimate the extent of under-reporting of cigarette production in Pakistan. The results confirm a significant level of under-reporting.[5]

This study differs from previous research in two ways. First, instead of annual data, it uses quarterly data to estimate the extent of under-reporting, thus allowing the estimates to correspond with the fiscal year (July-June) rather than the calendar year. Second, the previous research only included two large firms listed on the Pakistan Stock Exchange (PSE), whereas this study includes a third firm to cover all PSE-listed tobacco companies.[6] This will make it more feasible to compare officially reported large-scale manufacturing of cigarettes with estimated production.

Research context

Although the debate over the illicit trade of cigarettes in Pakistan's market is not new, it has gained momentum in recent years. For instance, Khan et al. (2021) analyzed cigarette packs collected from 10 cities in Pakistan to estimate the magnitude of illicit cigarettes. Their definition of illicit cigarettes includes those packs that lack any of the following: text and pictorial health warnings, underage sale prohibition warnings, retail price, manufacturer's name, and cigarette packs sold for less than the printed price. The analysis is based on a consumer survey and a waste and recycling store survey; both conducted between September 2019 and March 2020. According to the findings, one out of every six cigarette packs consumed in Pakistan could be termed illegal.

Another study (FFO, 2019) examined the magnitude of illicit cigarettes in Pakistan's capital, Islamabad, using the empty pack collection method. A cigarette pack was considered illegal if any of the following items were missing: text and pictorial health warnings, underage sale prohibition warnings, retail price, and the name and location of the manufacturer. According to the findings, non-duty paid cigarettes accounted for 5.6 percent of total trade, while illicit cigarettes accounted for 15 percent of total trade. They defined "non-duty paid cigarettes" as cigarettes meeting all legal criteria but were sold for less than the printed price in the retail market.

[5] The estimates of supply function based on monthly time series of production and prices of cigarettes indicate that the firms under-reported their production by 47 percent and 27 percent in fiscal years 2016-17 and 2017-18, respectively. Similarly, the results of econometric model based on financial panel data suggested that under-reporting was 39.5 percent and 21.5 percent in 2017 and 2018, respectively.

[6] There are three PSE-listed companies—Pakistan Tobacco Company (PTC), Phillips Morris Pakistan (PMP) and Khyber Tobacco Company (KTC). Euromonitor data shows that the market of PTC and PMP is 71.0 percent and 26.8 percent, respectively. All other companies (including KTC) have a share of 2.2 percent in total licit production in the country, except Azad Jammu and Kashmir.

CTFK (2019) also estimated the size of the illicit cigarette trade based on a survey of 2,111 retailers in ten cities across Pakistan. The data was collected in 2016, prior to the introduction of the third tier of FED. They examined 7,496 cigarette packs using a set of six indicators. Findings show that the size of illicit trade was around 9 percent; however, as there is no system of placing tax stamps on packs, it was nearly impossible to determine whether the company had actually paid the tax amount printed on the pack. Moreover, they discovered that 13.8 percent of consumers paid less than the minimum retail price notified by FBR, which indicates tax evasion.[7] In contrast, Oxford Economics (2018)[8] and Euromonitor-International (2021) claim a high magnitude of illicit cigarettes in Pakistan. According to them, illicit trade hovered around 31 billion to 33 billion sticks in 2017 and 2020.[9] In relative terms, it increased from 29.7 percent in 2017 to 36.4 percent in 2020.

In contrast with the survey-based studies mentioned above, Iqbal et al. (2020) used econometric techniques to estimate the size of under-reported cigarette production by the tobacco industry in Pakistan. The production function was estimated using annual financial data from firms, and the supply function was estimated using macro data on prices and production. According to the findings, under-reporting was 39.5 percent in 2017, which fell to 21.5 percent in 2018 after the introduction of the third tier and reduction of tax rates.

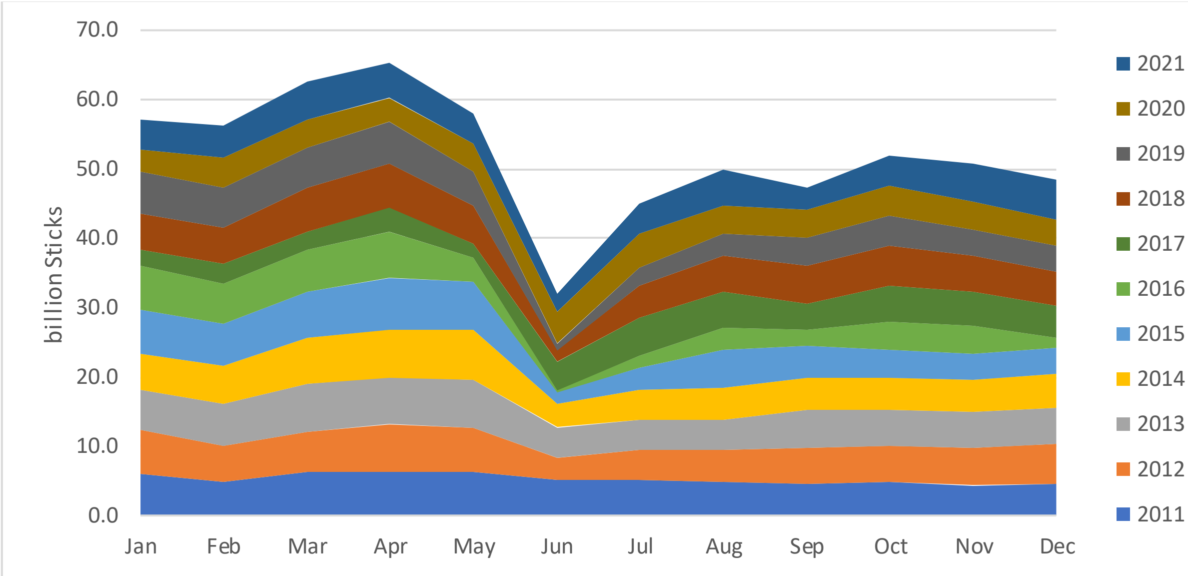
Moreover, the volume of undeclared cigarette production may vary depending on the expected changes in tax policy. Figure 1 depicts the declared monthly production of Pakistan's cigarette industry from 2011 to 2021. The data reveals an intriguing pattern related to fiscal policy announcements. The fiscal year in Pakistan runs from July to June. The federal budget and finance act, which may include amendments in tax rates or structure, is presented in the National Assembly by the government in the month of June. It is evident that declared production in the months preceding the budget (January to May) is significantly higher than the production reported in the months following the budget (June to December). This suggests that, in addition to under-reporting, the cigarette industry uses front-loading to avoid taxes.

Interestingly, the behavior appears to have changed over time with increased front-loading in recent years. The front-loading from 2013-14 to 2019-20 seems to be in response to the tax policy during this period; FED tax rates were raised every year except in 2017-18. Given the expectations of tax changes, the industry used front-loading to prevent or reduce the impact of the tax hikes. Furthermore, compared to the rest of the year, declared production in June is

[7] As per FBR selling cigarettes less than minimum price is illegal that can be linked to tax evasion.
[8] It is important to mention that the Oxford Economics prepared this report with the financial support of Phillip Morris International.
[9] As per the data of Pakistan Bureau of Statistics, documented production of cigarettes was 34 billion and 46 billion sticks in 2016-17 and 2019-20 (see Figure 1 for year-wise production).

significantly lower, reflecting the industry's intention to put pressure on the government to refrain from raising taxes.

Figure 1. Monthly cigarette production trends in Pakistan



Source: Quantum Index of Large Scale Manufacturing Industries, Pakistan Bureau of Statistics, https://www.pbs.gov.pk/qim

Based on the discussion above, it appears that Pakistan's market contains a sizable proportion of undocumented production of cigarettes, and the industry uses this to influence tobacco taxation policy. In this context, this study uses econometric techniques to estimate the size of the under-reported production of cigarettes. The scope of the research is limited to undeclared domestic production by the manufacturers from 2018-19 to 2020-21.

Research objective

The study's objective is to quantify the under-reporting of cigarette production in Pakistan and assess the extent of tax evasion by the cigarette manufacturing industry from 2018-19 to 2020-21. This is accomplished by analyzing the value of input supplies and output as reported in the financial statements of cigarette manufacturing firms. The analysis involves estimating the production function, converting the fitted value to the aggregated numbers of sticks, and comparing them with the reported sticks for computation of the magnitude of under-reported cigarettes. The analysis will help determine the revenue loss to the government due to under-reporting. The findings will also assist in estimating potential revenue gains as a result of the implementation of the Track & Trace System,[10] along with contributing to the development of

[10] With effect from July 01, 2022, the FBR has rolled out an electronic track and trace system across four sectors, including tobacco, cement, sugar and fertilizers. It involves affixation of tax tamps on various products, including cigarette packs. The system is being implemented by Pakistan Tobacco Company, Philip Morris Pakistan, and Khyber Tobacco Company. However, the effectiveness of this system is yet to be seen.

a case for a higher tax rate since the decrease in reported quantity does not correspond to a reduction in actual production.

Research limitations

The scope of the research is limited to the linkages between financial data on input supplies and the value of turnover. The empirical strategy relies heavily on production function estimation based on financial statements of three listed companies and assumes a constant share of factors of production over the estimated period. There may be other non-listed small-scale firms, particularly in Azad and Jammu Kashmir, which are not included in the analysis due to the unavailability of any information about these firms.[11] Moreover, the financial data used for the study is obtained from the financial statements of the PSE-listed firms that are publicly available online. It is beyond the scope of this research to cross-verify financial data from the firms.

[11] SPDC (2020) has shown that the AJK cigarette industry is not a big player in Pakistan's tobacco industry, since the share of AJK in relation to total production of cigarettes in Pakistan is only about one percent.

Literature Review

Iqbal et al. (2020) presented a brief review of relevant academic and grey literature divided into two research streams. The first set of literature focuses on methodologies used in tobacco-related research to estimate the magnitude of non-duty-paid cigarette sales and consumption. The second body of literature focuses on the econometric estimation of production functions. The following paragraphs provide a brief refresher by highlighting the key message from Iqbal et al. (2020), along with reviewing a couple of recent studies.

Research methodologies for estimating the illicit trade of cigarettes

Ross (2015) provided an excellent review of methodologies for estimating the magnitude of the illicit trade and tax gap by quantifying non-duty-paid cigarettes. She classified these methodologies into five categories; however, after weighing the benefits and drawbacks of different approaches in the context of Pakistan, the review (Iqbal et al., 2020) concluded that none of the methods fully corresponds to the scope of research, despite the fact that it can be linked to the category of empirical studies. Further, it concluded that a production function approach would be appropriate for estimating the extent of under-reporting for the study.

Methodologies for estimating the production function

The second set of literature in Iqbal et al. (2020) focuses on the econometric estimation of production functions and includes three studies. One of the studies (Ringstad, 1971) exemplifies the use of various statistical methods to estimate sector-specific CES (constant elasticity of substitution) production functions for mining and manufacturing industries in Norway. Another study (Ackerberg et al., 2015) proposes an alternative approach, arguing that some popular techniques for estimating production functions may suffer from a functional dependency problem. The third study (Grieco et al., 2016) proposes an alternative method to estimate production functions in the presence of input price dispersion when intermediate input quantities are not observed.

Moreover, Gandhi et al. (2020) noted that there is a growing interest in the literature in estimating gross output models that include intermediate inputs. They present new results for the non-parametric identification of gross output production functions in the presence of both flexible and non-flexible inputs using the model structure of the proxy variable approach. They used two commonly used firm-level production datasets to demonstrate that their non-parametric estimator provides reasonable estimates of production function elasticities.

Felipe et al. (2021) investigated the issue of estimating the production function using financial and physical data. Their main argument is that the model's functional form is unimportant because financial data correspond to production identity. As a result, regardless of functional forms or estimation techniques, results will lead to factor shares rather than production elasticities, as in the case of physical data. To support their argument, they reported Cobb-Douglas estimates of the production function using five estimation methods: ordinary least squares (OLS), least-squares dummy variables (LSDV), instrumental variables (IV), system generalized method of moments (GMM), and Levinsohn and Petrin (L-P). The study is based on a dataset that spans 473 industries over 54 years. They concluded that all of the above methods produced roughly the same factor shares; however, their simple OLS produced more plausible results. Furthermore, their findings indicate that theoretically more appropriate estimators for dealing with endogeneity, such as GMM or L-P, do not appear to produce better results.

The reviewed literature hence demonstrates a variety of approaches that can be used in the estimation of the production function. But, in line with the research scope, this study heavily relies on Felipe et al. (2021) to estimate production function from financial data.

Methodology

A three-step methodology is adopted to estimate the magnitude of under-reported production of cigarettes. The first step is to construct a database of available variables from the financial statements of three PSE-listed companies and other relevant secondary sources. The second step is to estimate a theoretically consistent production function. The third and final step is to compute the aggregate annual values of production, convert them into the number of cigarette sticks, and calculate the magnitude of under-reported cigarettes by taking the difference between the number of reported and estimated sticks.

Step-1: Data sources and construction of variables for estimation

The database of panel data is constructed by using firm-level information from quarterly financial data of all PSE-listed tobacco companies (PTC, KTC, and PMP).[12] Financial records of the firms provide information about manufacturing costs and sales. Except for KTC, consistent data was available in the annual reports of firms from the first quarter of 2012 to the second quarter of 2021. Quarterly statistics for KTC were available from the start of the third quarter of 2012. The data on net turnover (turnover after excluding taxes) is used as a proxy for the value of output. Similarly, the wage bill is also obtained from the financial statements of the firms, while profit is computed after subtracting all input costs from the turnover.

Aside from the firm-level financial data, time-series data on cigarette prices, cigarette price indexes, aggregate cigarette production, and general CPI are obtained from the different documents/website of the Pakistan Bureau of Statistics (PBS). For instance, the aggregated output (aggregate cigarette production) is gathered from the quantum index of large-scale manufacturing industries.[13] The monthly price bulletin and its annexures are used to obtain all price variables and their corresponding price indices.[14]

The following variables are constructed to estimate the number of under-reported sticks and the production function. Firstly, nominal variables are constructed to be used as dependent and independent variables in the production function. The value of "output" is the proxy for declared production and is computed by gross turnover netted out by indirect taxes and adjusted

12 The data is obtained from the following websites last accessed on February 28, 2021. For quarterly financial statement of PTC: https://www.ptc.com.pk/group/sites/PAK_AMPC26.nsf/vwPagesWebLive/DOANGKM8, for quarterly financial statement of PM: https://www.pmi.com/investor-relations/reports-filings, and for quarterly financial statement of KTC http://www.khybertobacco.com/financial-statements/.

13 The quantum index of large-scale manufacturing industries is available on the website https://www.pbs.gov.pk/qim (last accessed on February 28, 2021)

14 For recent months, both are available at https://www.pbs.gov.pk/cpi-nb. Earlier data is obtained from the monthly bulletin of Pakistan Bureau of Statistics.

for inventories. The wage bill consists of production-related salaries, while the cost of intermediate inputs is the sum of the costs of raw materials, fuel and power, storage, and other costs. Profit is calculated by subtracting the cost of intermediate inputs and the wage bill from the net value of output.

All these variables are deflated and converted into real values using the cigarette price index (CPI-cigarette). There are two reasons for using the CPI-cigarette instead of the CPI-general. First, the cigarette industry's output, raw materials, storage, salary bill, and other costs are all inextricably linked and, to some extent, manipulated by the industry. While the CPI-general and the CPI-cigarette have a correlation of nearly 0.9, we believe the CPI-cigarette is more relevant and well-suited for deflating these variables. Second, to ensure that the production function identity holds, we want to deflate all variables with the same price index. Again, CPI-cigarette is a better option.

Step-2: Estimation of production function

Theoretically, the production function estimates the relationship between inputs and outputs by using various functional forms. In other words, a production function is a description of a production technology that relates the physical output of a production process to the inputs or factors of production.

In a panel data setting, a general representation of the production function is as follows:

$$Q_{it} = f(A_{it}, L_{it}, K_{it}, Z_{it}) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where Q, L, K, and Z represent physical output, number of workers, physical capital stock, and intermediate consumption, respectively; and A is some measure of technology, and subscripts i and t refer to firms and time. In a physical production function, the output must be determined by accurately measured flow of services from labour (number of workers) and capital (number of identical machines) and the rate of material utilization (kilowatts of electricity, or magnitude of raw materials). As argued by Felipe et al. (2021), estimating such a production function is practically nearly impossible due to the data requirements.

In practice, data for the monetary value of output, capital, labour wage bill, and intermediate input cost are available. The transition from physical to monetary data results in the following equation of output.

$$Y_{it} = W_{it} + P_{it} + Z_{it} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

where in equation 2, real gross output (Y) is the sum of real wage bill (W), real total profits (P), and real cost of intermediate inputs (Z).

Dividing equation 2 by Y_{it} yields the following equation.

$$1 = W_{it} / Y_{it} + P_{it} / Y_{it} + Z_{it} / Y_{it} \quad \ldots.. (3)$$

We assume a, b, and c are the factor share of labour, capital and intermediate inputs and have following values.

$$a_{it} = W_{it} / Y_{it},\ b_{it} = P_{it} / Y_{it}, \text{ and } c_{it} = Z_{it} / Y_{it}$$

Equation 3 can be rewritten as:

$$1 = a_{it} + b_{it} + c_{it} \quad \ldots.. (4a)$$

$$c_{it} = 1 - a_{it} - b_{it} \quad \ldots.. (4b)$$

As Felipe et al. (2021) pointed out, equation 2 holds true at all levels of aggregation, from firm to economy-wide. It is also compatible with any type of market and degree of returns to scale. It holds even when there is no well-defined production function because there are no economic or other assumptions about factor markets or the degree of scale returns involved.

The total differentiation of equation 2 with respect to time yields the following equation.

$$dY_{it} = dW_{it} + dP_{it} + dZ_{it} \quad \ldots.. (5)$$

Dividing both sides by Y_{it} and arranging it in growth rates.

$$\frac{dY_{it}}{Y_{it}} = \frac{W_{it}}{Y_{it}}\frac{dW_{it}}{W_{it}} + \frac{P_{it}}{Y_{it}}\frac{dP_{it}}{P_{it}} + \frac{Z_{it}}{Y_{it}}\frac{dZ_{it}}{Z_{it}} \quad \ldots.. (6)$$

Substituting the value of factor shares from equation 4a and 4b

$$\frac{dY_{it}}{Y_{it}} = a_{it}\frac{dW_{it}}{W_{it}} + b_{it}\frac{dP_{it}}{P_{it}} + (1 - a_{it} - b_{it})\frac{dZ_{it}}{Z_{it}} \quad \ldots.. (7)$$

Assuming factor shares are constant across time.[15]

$$\frac{dY_{it}}{Y_{it}} = a\frac{dW_{it}}{W_{it}} + b\frac{dP_{it}}{P_{it}} + (1-a-b)\frac{dZ_{it}}{Z_{it}} \qquad \ldots\ldots (8)$$

Integrating equation 8 yields.

$$Ln(Y_{it}) = d + aln(W_{it}) + bln(P_{it}) + (1-a-b)\ln(Z_{it}) \qquad \ldots\ldots (9)$$

where d is the constant of integration. Equation 9 is the desired equation that links financial values of inputs to the value of output and used for estimation.

Step-3: Computation of annual output and under-reported cigarettes

Fitted values obtained from equation 9 is used to estimate the aggregated production.

$$\widehat{Y_{it}} = exp(\widehat{lnY_{it}}) \qquad \ldots\ldots (10)$$

Since, the estimated is real turnover, to convert the series into nominal values it is multiplied with the consumer price index of cigarettes. Mathematically,

$$\widehat{Y_{it}} = exp(\widehat{lnY_{it}}) * CPI_{cig}/100 \qquad \ldots\ldots (11)$$

The quarterly aggregates value of output is obtained by using the following formula.

$$\widehat{Y_t} = \sum_{i=1}^{3}\widehat{Y_{it}} \qquad \ldots\ldots(12)$$

After having the aggregated value, the annual value for each fiscal year is simply the sum of four quarterly values starting from July-September and ending in March-June for the corresponding year.

$$\widehat{Y_T} = \sum_{t=July-Sep}^{Mar-June}\widehat{Y_t} \qquad \ldots\ldots(13)$$

where T indicates fiscal year. Equation 13 is used to compute the aggregate nominal value of output (turnover) for the fiscal year 2018-19, to 2020-21.

After having the value of output for each fiscal year, the next step is to compute the number of sticks and compare them with reported data to compute the magnitude of under-reporting.

[15] This is the only assumption applied in the derivation.

While the firm-wise number of stick data is not consistently available in the quarterly report, the average producer price was computed by dividing the reported output value (turnover) by the number of sticks for each fiscal year.

$$PP_T = \frac{Y_T}{N_T} \qquad \text{..... (14)}$$

where PP_T indicates net of taxes producer price of cigarettes (PKR per stick). This price is used to estimate the number of estimated cigarettes for each fiscal year.

$$\widehat{N_T} = \frac{\widehat{Y_T}}{PP_T} \qquad \text{..... (14)}$$

The under/over reporting is the difference of reported and estimated number of sticks.

$$\text{Difference} = \widehat{N_T} - N_T \qquad \text{..... (16)}$$

To sum up, the magnitude of under-reported production of cigarettes is calculated using the fitted series from the estimates of the production function. The out-of-sample forecasting method is used to estimate the log of real turnover for the period beginning the third quarter of 2018 and ending the second quarter of 2021 for each firm. Afterward, using equation 11, the turnover is converted into nominal turnover. Following that, the combined quarterly turnover is estimated using equation 12. Finally, using equation 13, these quarterly turnovers are added to obtain the total turnover for each fiscal year (July-June) from 2018-19 to 2020-21.

After calculating the output value for each fiscal year, the next step is to compute the number of sticks and compare them to the reported data to determine the magnitude of under-reporting. As shown in equation 14, the producer price is calculated by dividing the actual net turnover by the officially reported number of sticks. As reflected in equation 15, these prices are used to calculate the estimated number of cigarettes. Finally, the magnitude of unreported cigarettes is determined by comparing estimated and reported cigarette stick counts.

Results

This section is divided into two parts: descriptive analysis and empirical results. The descriptive analysis covers firm-specific trends in market structure, turnover, and cost structure. The second part presents the results of econometric estimation of the production function and the estimates of under-reporting of cigarette production.

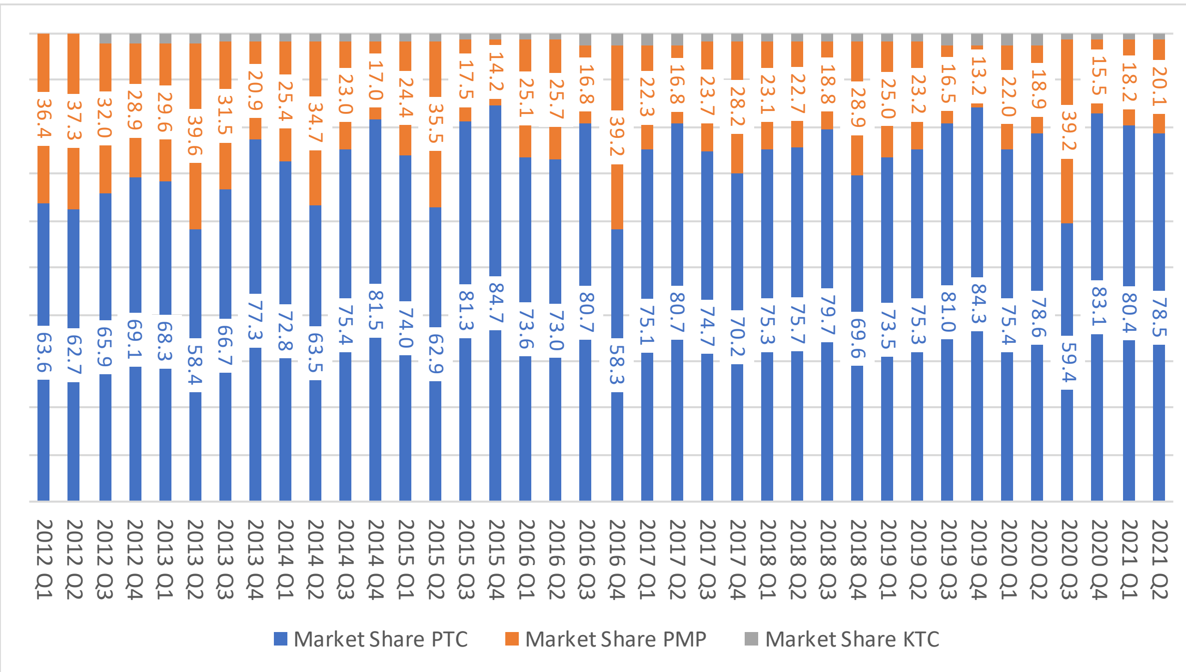
Descriptive analysis

Even though the cigarette industry is one of Pakistan's large-scale manufacturing industries, only three companies are listed on the Karachi Stock Exchange—PTC, PMP, and KTC. This section summarizes the key indicators extracted from the annual/quarterly reports of these companies. It focuses primarily on trends in their inputs, value-added, and outputs over a nearly ten-year period.

Market structure

Figure 2 presents the market share of each PSE-listed firm in the cigarette industry. Three significant observations emerge from the data. First, PTC is the leader in cigarette manufacturing, followed by PMP and KTC. Second, PTC's market share has increased over time while PMP's share decreased. For instance, PTC's average share was 65 percent in 2012, which rose to around 80 percent during the first two quarters of 2021. In comparison, the average market share of PMP fell from 33 percent to about 19 percent during the same period. The KTC has a relatively small market share hovering around two percent.

Figure 2: Trend in market share based on the reported turnover of the listed companies (%)



Source: Authors' estimates based on Quarterly Financial Statements of the tobacco companies

Finally, the data show quarterly variations in market share during the majority of the years, with PTC having the highest share in the fourth quarter and PMP having the highest share in the second quarter of each year.

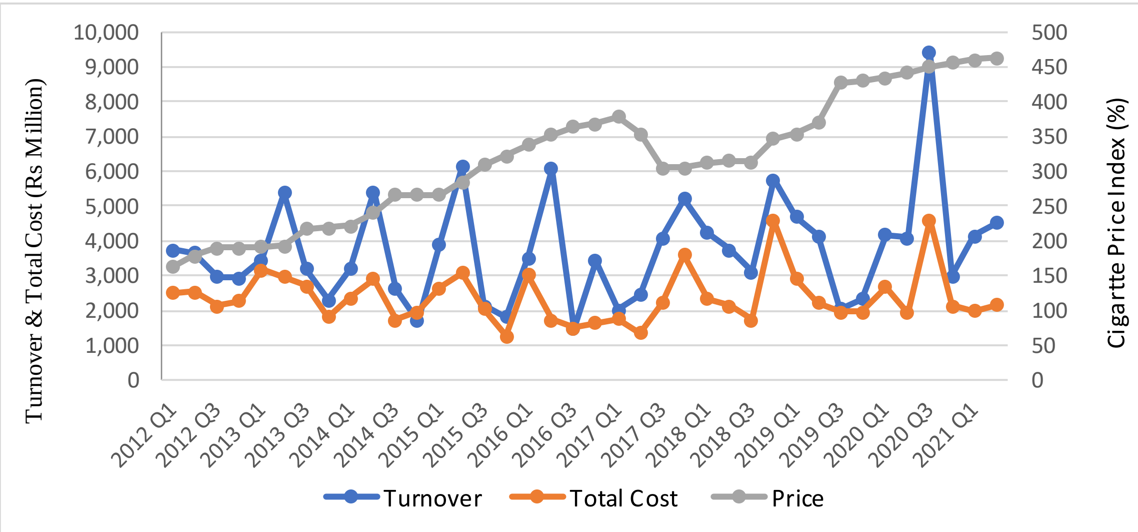
Trends in financial data

Figures 3a, 3b, and 3c show the quarterly nominal turnover and total cost for PMP, PTC, and KTC, respectively, as well as the market price index for cigarettes. The quarterly turnover includes net inventories (current stock minus previous stock) but excludes indirect taxes. The price index contains both taxes and producer costs.

Phillip Morris Pakistan

The trend in financial data of PMP highlights three messages. Firstly, there is consistent seasonal variation in cigarette production as the highest production is reported in the first quarter and the lowest in the third quarter of each year until 2016. Following that, the highest cigarette output shifted to the fourth quarter, and overall cigarette manufacturing peaked in the third quarter of 2020. Secondly, total costs have historically been highly correlated with turnover except for the last two quarters when turnover has increased while costs have remained stable. Finally, there is absolutely no correlation between the market price of cigarettes and the total cost.

Figure 3a: Total cost and turnover – Phillip Morris Pakistan

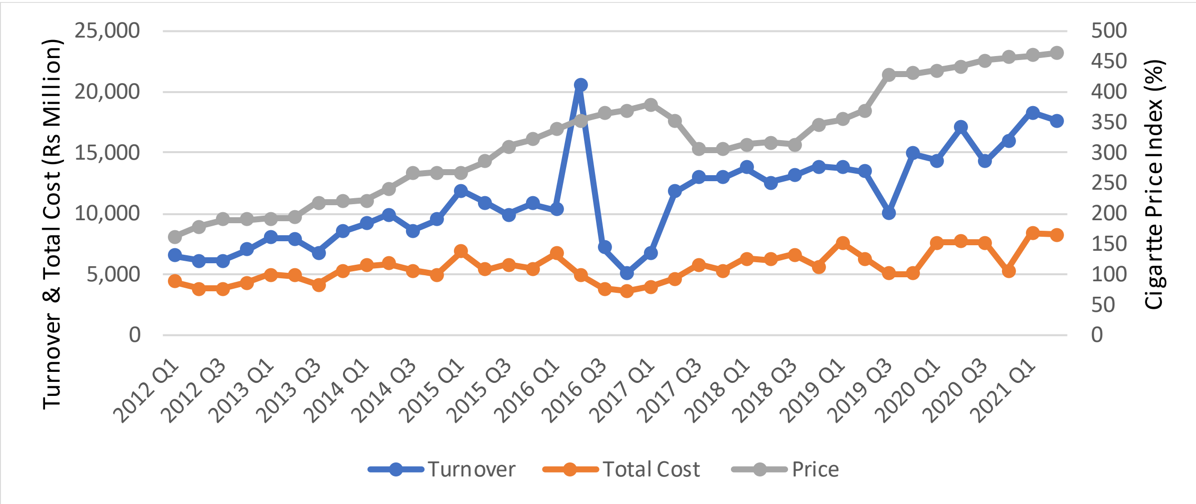


Source: Quarterly Financial Statements of Phillip Morris Pakistan and Pakistan Bureau of Statistics

Pakistan Tobacco Company

It was expected that PTC, being the market leader, would be driving the market price. However, the trend of total cost and market price presented in Figure 3b does not show similar fluctuations and reflects a weak correlation of 0.53, which is much lower compared to PTC's market share. Instead of the total cost, the market price of cigarettes appears to be sensitive to the FED rate. Furthermore, turnover has no clear seasonal pattern (Figure 3b). For the majority of the quarters, the turnover is on the rise. The only exception was the second quarter of 2016 when an increase in cigarette prices due to an increase in the FED rate resulted in a sharp drop in the reported turnover. Finally, compared to the turnover, the total cost has a more stable trend with fewer fluctuations.

Figure 3b: Total cost and turnover – Pakistan Tobacco Company

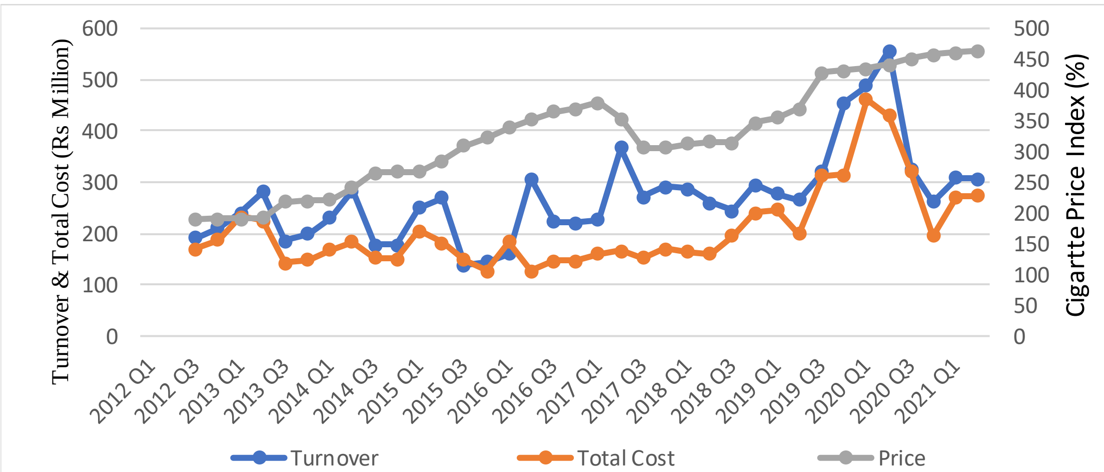


Source: Quarterly Financial Statements of Pakistan Tobacco Company and Pakistan Bureau of Statistics

Khyber Tobacco Company

As shown in Figure 3c, KTC's turnover has a strong seasonal pattern with a peak in the second quarter most of the time. Like PTC, KTC has a correlation of 0.55 between market price and turnover. Also, the turnover has a strong association with total cost.

Figure 3c: Total cost and turnover – Khyber Tobacco Company



Source: Quarterly Financial Statements of Khyber Tobacco Company and Pakistan Bureau of Statistics

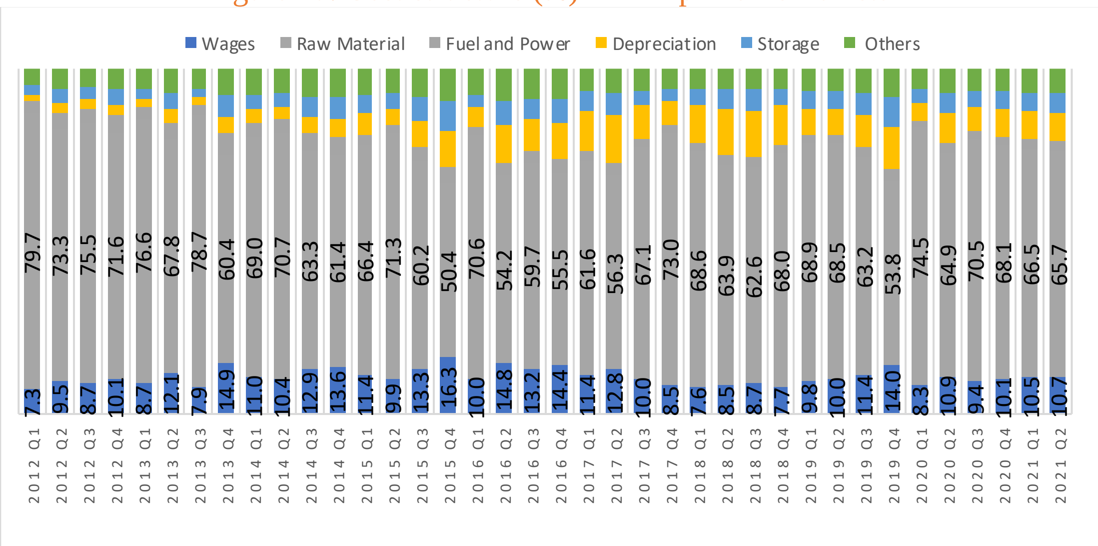
Firm-wise cost structure

Figures 4a, 4b, and 4c present quarterly shares of different cost components for PMP, PTC, and KTC, respectively. The cost components include wages, raw material costs, fuel and power costs, depreciation, storage, and other costs. These shares are based on nominal values.

Phillip Morris Pakistan

PMP data shows that raw materials and wage bills are the two most significant cost components, accounting for at least three-fourths of the overall cost (Figure 4a). Raw material costs follow a seasonal trend, with the highest share in the first or fourth quarter of the calendar year. Raw material had an average share of 75 percent in 2012, which gradually fell to 60 percent in 2016 and rebounded to nearly 70 percent in 2020.

Figure 4a: Cost structure (%) – Phillip Morris Pakistan

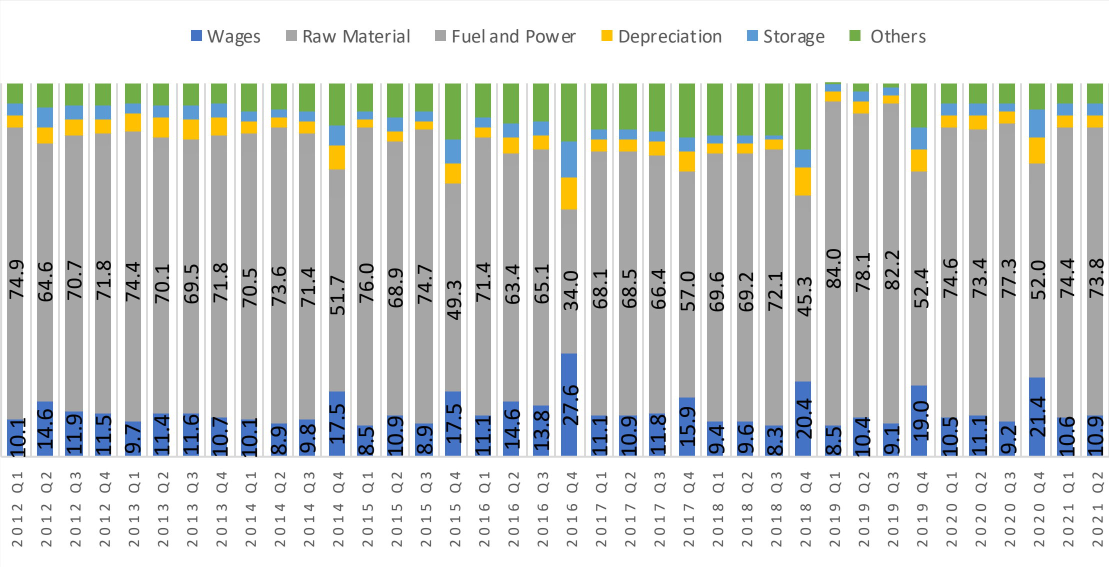


Source: Quarterly Financial Statements of Phillip Morris Pakistan

Pakistan Tobacco Company

As far as the trend in cost components of PTC is concerned, the data reveals that the cost of raw materials has a relatively stable share hovering around 70 percent, with the exception of 2016, when the share of raw materials was less than 60 percent on average (Figure 4b). The raw material share also varies seasonally—the lowest share being in the fourth quarter most of the time. The fourth quarter of 2016 shows an unusual lower share of 34 percent. An increase in the share of wages and other costs usually occurs when the share of raw materials falls.

Figure 4b: Cost structure (%) – Pakistan Tobacco Company

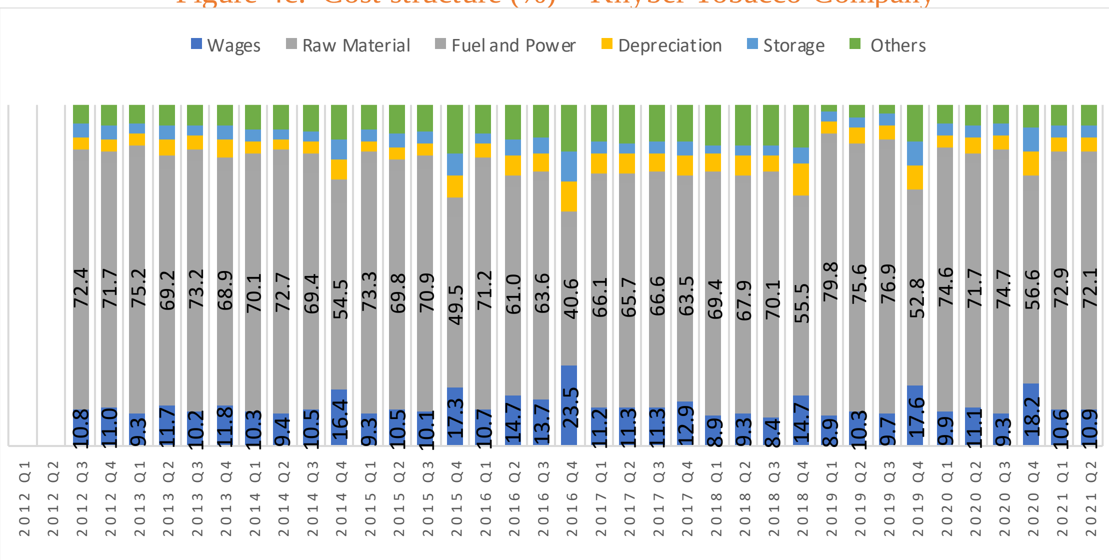


Source: Quarterly Financial Statements of Pakistan Tobacco Company

Khyber Tobacco Company

In the case of KTC, the cost of raw materials has a seasonal pattern, with the lowest share in the fourth quarter of the calendar year. Particularly, the fourth quarter of 2016 shows an unusually low share of 40 percent. When the share of raw materials falls, the share of wages and other costs often rises.

Figure 4c: Cost structure (%) – Khyber Tobacco Company



Source: Quarterly Financial Statements of Khyber Tobacco Company

Empirical Results

Step-1: Construction of variables production function estimation

The following variables were constructed in the first step: log of real output (lrturnover), log of real profit (lrprofit), log of real wages (lrwages), and log of real intermediate inputs (lrin_inputs). Summary statistics are provided in Annex-A. The summary statistics show a total of 112 quarterly data observations, divided into three panels representing PMP, PTC, and KTC. PMP and PTC both have 38 quarterly data points spanning all four series from the first quarter of 2012 to the second quarter of 2021. KTC has 36 quarterly data points ranging from the third quarter of 2012 to the second quarter of 2021.

Step-2: Production function estimation

The empirical estimation of the production function is divided into three steps. These include testing stationarity and order of cointegration, model estimation, and sensitivity and robustness analysis by estimating four models and comparing fitted and actual series to select the best model.

Testing stationarity and order of cointegration

In the first step, two stationarity tests, "Levin-Lin-Chu unit-root test" and "Im-Pesaran-Shin unit-root test," are used to determine the order of integration. The results of both tests are summarized in Annex-B. Both tests confirmed that all four series are stationary, with variables measured at levels. Following the stationarity tests, three cointegration tests, the "Kao test," "Pedroni test," and "Westerlund test," are used to confirm the order of integration in panel data sets. The results of all three tests are summarized in Annex-B. All three tests confirmed that the dependent variable lrturnover cointegrated with the explanatory variables lrprofit, lrwages, and lrin_inputs at the level.

Estimation of the base model

In the third step, equation 9 is transformed in the following for estimation in the panel data set.

$$\text{lrturnover}_{it} = \pi_0 + \alpha\text{lrwages} + \beta\text{lrprofit}_{it} + \gamma\text{lrin\_inputs}_{it} + v_{it} \qquad \ldots.. (17)$$

The model in equation 17 is a transformation of equation 9 in a Common Effect Model or Pooled Least Square (PLS), where v is the uncorrelated error term. Equation 17 is estimated using four methods: common effects, fixed effects, least-square dummy variables (LSDV), and random effects procedures. Following inspiration from the literature (Felipe et al., 2021; Kumbhakar et al., 2015; Nguyen et al., 2021; Schmidt & Sickles, 1984), these estimation techniques were used to select the best model for estimating under-reported cigarettes.

Table 1 summarizes the results of four estimated models, while Annex-C contains the full estimation outputs. The qdate (quarterly dates) is added in all models to capture the impact of the time trend. The estimation of the fixed effects model is the first step in model selection. Except for the constant and log of real wages, the estimated results show that all variables are statistically significant. However, the F test that all u_i=0 was insignificant, arguing for a common effects model rather than a fixed-effects model.

In the next step, the LSDV model is estimated to validate the results of whether the firm-wide impact is significant or not. According to the findings, the estimated values of firm-specific dummy variables are statistically insignificant. These findings confirmed that the fixed effects model is not appropriate for the dataset of cigarette manufacturing firms.

Table 1: Summary of Estimated Results (Dependent variable is lrturnover)

lrturnover	Fixed Effects		LSDV		Random Effects		Common Effects	
	Coef.	t -stats	Coef.	t -stats	Coef.	t -stats	Coef.	t -stats
lrprofit	0.0644	7.05	0.0644	7.05	0.0661	7.46	0.0661	7.46
lrwages	0.1504	1.48	0.1504	1.48	0.2305	3.17	0.2305	3.17
lrint_inputs	0.7497	9.58	0.7497	9.58	0.7522	10.31	0.7522	10.31
PMP			0.1696	0.67				
ptc			0.3172	0.95				
ktc			0.0000	(omitted)				
qdate	0.0117	2.94	0.0117	2.94	0.0141	5.01	0.0141	5.01
_cons	-1.182	-0.63	-1.348	-0.80	-2.632	-3.67	-2.632	-3.67
R2 (adj. or overall		0.9890		0.9887		0.9892		0.9885
F/Wald chi2		74.37		1097.32		6477.48		1619.37
For fixed effects: F test that all u_i=0: F(2, 69) = 1.57						Prob > F = 0.2162		

Source: Authors' estimates

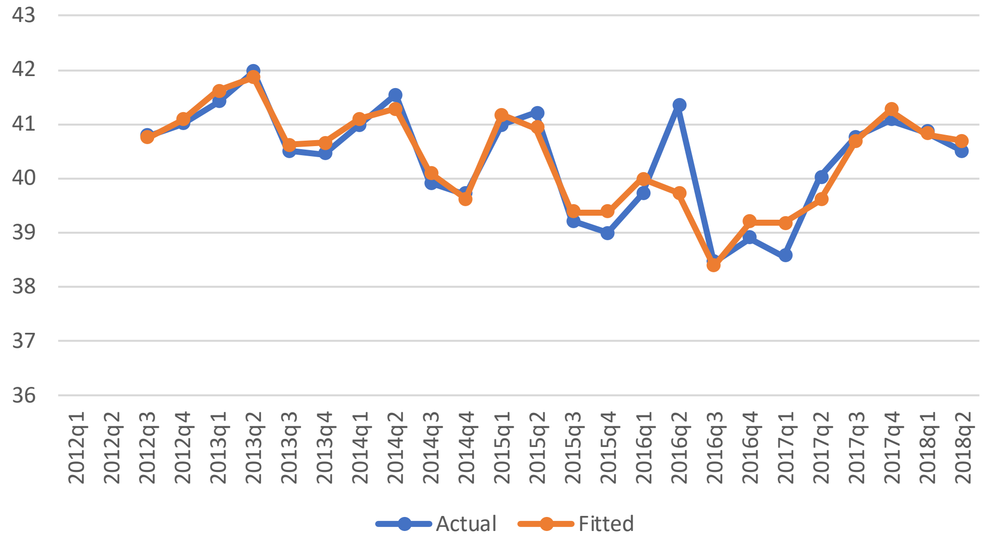
The random effects and common effects models were then estimated. The magnitudes of all estimated variables are the same in both models, and all variables are statistically significant with relatively high adjusted/overall R^2. Logs of real explanatory variables wages, capital, and intermediate inputs are all positive. Subsequently, the Hausman test was applied to validate the selection of a random-effects model over the fixed effects model. While the magnitudes of the random and common effects models are the same, the common effects model was used to estimate the extent of under-reporting.

Post-estimation comparison of fitted and actual series

Figure 5 shows the combined actual and fitted logs of three companies' real turnover. The trend indicates that until the third quarter of 2015, the fitted series closely matched the actual series.

The second quarter of 2016 proved to be an outlier. After the second quarter of 2017, the fitted series closely matches the actual series. The convergence of the fitted series with the actual series over most of the period suggests that it is suitable for ex-post forecasting of real and nominal turnover.

Figure 5: Three companies combined actual and fitted logs of real turnover



Sensitivity and robustness

Seasonality and front-loading are evident in the quarterly data. Apart from quarterly variations, the cigarette firms reported record production in the second quarter of 2016, which coincides with a change in tax policy, where FED rates on cigarettes were increased in the Federal Budget 2016-17.[16] It appears that the firms over-reported the production in the second quarter of 2016 and under-reported in the rest of 2016 to minimize the impact of tax rate increases and put pressure on the government to restrain tax increases.

A sensitivity analysis is undertaken by generating three different models to examine the influence of seasonal variations and substantial front-loading in the second quarter of 2016 on the empirical estimates. The first model includes all of the variables from the basic model as well as three quarterly dummies to account for seasonality. These dummies contain one for the respective quarters and zero for the rest of the time periods. Along with the base model, the second model included an outlier dummy variable with a value of one for the second quarter of three companies and zero for the remaining periods. Finally, the third model includes three quarterly dummy variables as well as an outlier dummy to estimate the combined impact.

[16] FED rates for lower and upper tiers were increased by 16.1 percent and 17.4 percent, respectively.

Table 2 summarizes the findings of three estimated models, whereas Annex-D gives the full set of estimation results. Each of the three quarterly dummies is statistically significant and has a negative sign. This confirms that the firms produce more in the second quarter than in the other quarters. The statistical significance of the other estimated coefficients was not affected by the insertion of quarterly dummies. The magnitude of the coefficients, however, varies slightly. Profit and input factor shares fell to 5.6 percent and 73 percent, respectively, while the factor share of salaries increased to 26.4 percent. The modified R-square suggests that the model's overall significance has improved marginally.

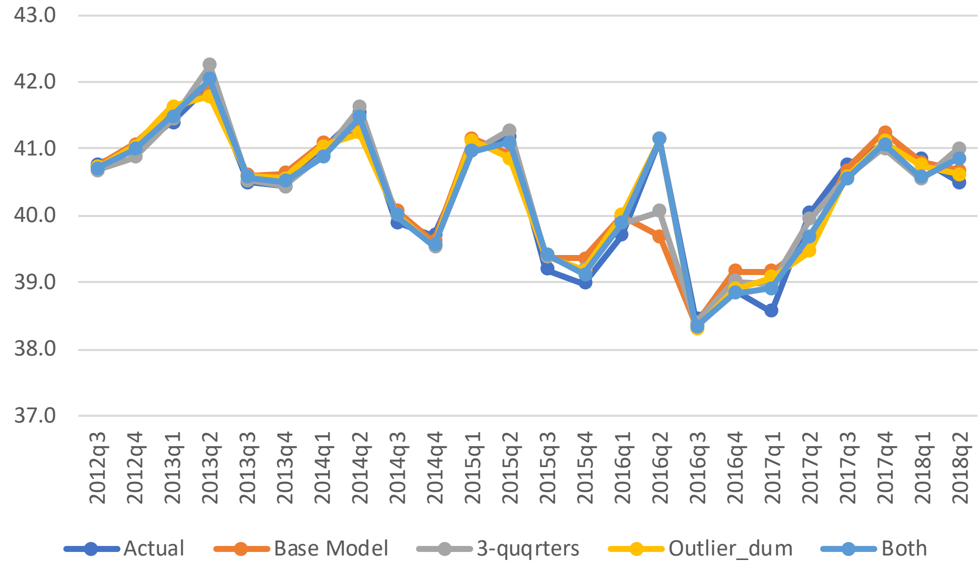
Table 2: Summary of Estimated Results (Dependent variable is lrturnover)

lrturnover	Qarterly Dummies		Outlier Dummy		Both Dummies	
	Coef.	t -statistics	Coef.	t -statistics	Coef.	t -statistics
lrprofit	0.0563	*7.20*	0.0566	*8.59*	0.0524	*8.25*
lrwages	0.2637	*3.44*	0.1506	*2.77*	0.1507	*2.33*
lrint_inputs	0.7301	*9.51*	0.8454	*15.43*	0.8505	*13.04*
qdate	0.0135	*5.60*	0.0132	*6.38*	0.0132	*6.77*
out_dum			0.5510	*7.11*	0.4758	*6.08*
q1	-0.1879	*-3.88*			-0.1385	*-3.46*
q3	-0.1529	*-3.09*			-0.0894	*-2.16*
q4	-0.2008	*-3.73*			-0.0985	*-2.11*
_cons	-2.3192	*-3.69*	-2.6760	*-5.09*	-2.6086	*-5.11*
	Adj. R2	0.9917	Adj. R2	0.9938	Adj. R2	0.9946
	F(7,68) =	1,284	F(5,70) =	2,419	F(8,67) =	1,722

The outlier dummy is also statistically significant in the second model. As expected, it exhibits a positive sign showing the highest level of front-loading or over-reporting of production in the second quarter of 2016. Even after accounting for quarterly volatility, the outlier dummy has a statistically significant non-zero value, according to the last model.

Figure 6 shows the combined actual and fitted logs of three companies' real turnover. The trend indicates that until the third quarter of 2015, all fitted series closely matched the actual series. However, in the second quarter of 2016, only two models with the outlier dummy moved close together with the actual series. After the second quarter of 2017, the fitted series began to closely match the actual series. The fitted series' convergence with the actual series for most of the period implies that all these models can be used for ex-post real and nominal turnover predictions.

Figure 6: Three companies combined actual and fitted logs of real turnover



Step-3: Estimation of under-reported production

The magnitude of under-reported production of cigarettes is calculated using the estimated fitted series from the common effects model having both quarterly and firm-wise dummy variables. The out-of-sample forecasting method is used to estimate the log of real turnover for the period beginning in the third quarter of 2018 and ending in the second quarter of 2021 for each firm. Afterward, using equation 11, these turnovers are converted to nominal turnover. Following that, the combined quarterly turnover is estimated (equation 12). Finally, using equation 13, these quarterly turnovers are added to obtain the total turnover for each fiscal year (July-June) from 2018-19 to 2020-21.

After computing the output value for each fiscal year, the next step is calculating the number of sticks and comparing those with the reported data to determine the magnitude of under-reporting. As reflected in equation 14, the producer price is calculated by dividing the actual net turnover by the officially reported number of sticks. These prices are then used to calculate the estimated number of cigarettes (equation 15).

Table 3 presents the estimated volume of unreported production of cigarettes from 2018-19 to 2020-21. The estimates reveal that 4.73 billion cigarette sticks were under-reported in 2018-19. The volume of under-reporting declined to 3.65 billion sticks and bounced back to 4.46 billion sticks in 2020-21. In relative terms, the production of cigarettes was under-reported by 7.8 percent, 7.9 percent, and 8.6 percent in 2018-19, 2019-20, and 2020-21, respectively.

Under-reporting of cigarette production has profound implications for tax revenue in Pakistan. Altogether, the loss of FED revenue owing to unreported production is estimated to be PKR 23.5 billion during the three years. When including GST revenue, it becomes PKR 29.5 billion.

Table 3: The estimated extent of under-reporting

Year	Production		Under-reporting		Revenue Loss		
	Declared*	Estimated**			FED	GST	Total
	Million sticks		Million sticks	(%)	PKR billion		
2018-19	60,729	65,462	4,733	7.8	7,027	1,801	8,828
2019-20	46,085	49,736	3,651	7.9	7,165	1,852	9,017
2020-21	51,554	56,011	4,457	8.6	9,283	2,414	11,697
Aggregate	158,368	171,208	12,840	8.1	23,475	6,067	29,542

Source: * Declared output from Monthly Bulletin of Statistics, Pakistan Bureau of Statistics, various issues

** Estimated output authors estimate based on the estimated production function

Conclusions

Pakistan's cigarette manufacturing industry is an example of an imperfect market, with only three firms dominating total domestic production. Despite the small number of large manufacturers, the FBR relies on manufacturers' self-declaration of production to determine their tax liability. Unmonitored self-declaration provides opportunity for tax evasion and may lead to inefficiencies in the tax collection system.

This study aims to assess the extent of the under-reporting of domestic cigarette production in Pakistan by estimating the potential levels of output of the cigarette industry. The methodology is based on estimating the production function using quarterly financial panel data from three PSE-listed firms. A theoretically consistent model is used for estimation, assuming a constant factor share in production during the estimation period. This is a plausible assumption because it is unlikely that production technology would change dramatically in such a short period.

In absolute terms, the under-reporting of production was highest in 2018-19 and lowest in 2019-20. In relative terms, according to estimates, the extent of under-reporting was at its lowest in 2018-19 and gradually increased in subsequent years. Altogether, under-reporting of 12.8 billion sticks led to a revenue loss of Rs 29.5 billion in the three years of analysis. Interestingly, printed retail prices and FED rates did not increase during the last two years, implying that the gradual increase in relative under-reporting may be due to an attempt by the firms to cover the increases in variable costs during and after the pandemic. This supports the argument that self-declaration encourages those who want to evade tax through under-reporting of production.

Policy implications

Pakistan's large fiscal imbalances necessitate increased tax revenues, and tobacco taxation provides a window of opportunities in this regard. Improvements in tax administration aimed at tapping the full tax potential of the tobacco industry would result in a significant increase in revenues, as well as an increase in input costs for manufacturers, which would likely increase prices and lead to lower consumption. The following policy measures are recommended:

- *Integrate the FED to multi-stage taxes:* The FED is collected at the factory level on declared production, providing an incentive to under-report production. The FED and GST should be collected at three stages: at the factory, distribution, and wholesaler/retailer.
- *System for Electronic Monitoring of Production:* FBR has recently initiated the Track and Trace System (TTS), which can help enhance revenue, reduce counterfeiting, curb illicit

trade, and avoid front-loading. Since front-loading is a common practice in the cigarette industry, TTS can be used to discourage it by tying up the validity of tax stamps to one fiscal year, allowing the inventory of cigarettes with old stamps to be adjusted according to any change in the tax rate.

- *Link financial data to production:* The analysis demonstrates how to monitor tax evasion by analyzing company financial data. Such data mining will help the government monitor and address potential tax evasion and aid in developing a strong tax collection mechanism in the future.

- *Reduce stockpiling and front-loading:* The analysis reveals seasonality in reported production. Generally, production is high in the months preceding the federal budget announcement. This is primarily due to tax policy uncertainties and a lack of clear policy instruction about stockpiling. Anti-stockpiling measures and a medium-term tax policy guideline are needed to avoid front-loading.

References

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, *83*(6), 2411-2451.

Chaloupka, F. J., Powell, L. M., & Warner, K. E. (2019). The Use of Excise Taxes to Reduce Tobacco, Alcohol, and Sugary Beverage Consumption. *Annual Review of Public Health*, *40*(1), 187-201. https://doi.org/10.1146/annurev-publhealth-040218-043816

Chaloupka, F. J., Yurekli, A., & Fong, G. T. (2012). Tobacco taxes as a tobacco control strategy. *Tobacco control*, *21*(2), 172-180. https://doi.org/10.1136/tobaccocontrol-2011-050417

CTFK. (2019). *Study to assess the volume of illicit cigarette brands in Pakistan*

Euromonitor-International. (2021). *Passport: Tobacco in Pakistan*.

Felipe, J., McCombie, J., Mehta, A., & Bajaro, D. F. (2021). Production Function Estimation: Biased Coefficients and Endogenous Regressors, or a Case of Collective Amnesia? *Levy Economics Institute, Working Papers Series*.

FFO. (2019). *Taxation – the best way to cut tobacco use*.

Gandhi, A., Navarro, S., & Rivers, D. A. (2020). On the identification of gross output production functions. *Journal of Political Economy*, *128*(8), 2973-3016.

Grieco, P. L., Li, S., & Zhang, H. (2016). Production function estimation with unobserved input price dispersion. *International Economic Review*, *57*(2), 665-690.

Iqbal, M. A., Sabir, M., Saleem, W., Ali, A., & , & Aamir, N. (2020). *Quantifying the Potential Tax Base of the Cigarette Industry in Pakistan*. https://www.tobacconomics.org/files/research/569/Tobacco-Under-reporting-study-1.pdf

Khan, A., Dobbie, F., Siddiqi, K., Ansaari, S., Abdullah, S., Iqbal, R., Khan, Z., Sohail, S., Kanaan, M., & Huque, R. (2021). Illicit cigarette trade in the cities of Pakistan: comparing findings between the consumer and waste recycle store surveys. *Tobacco control*.

Kumbhakar, S. C., Wang, H., & Horncastle, A. P. (2015). *A practitioner's guide to stochastic frontier analysis using Stata*. Cambridge University Press.

Nguyen, B. H., Sickles, R. C., & Zelenyuk, V. (2021). *Efficiency Analysis with Stochastic Frontier Models using Popular Statistical Softwares*.

Oxford Economics. (2018). *Asia illicit tobacco indicator 2017: Pakistan*. https://illicittobacco.oxfordeconomics.com/media/OXFO5877_Pakistan_2018_Report.pdf

Ringstad, V. (1971). *Estimating production functions and technical change from micro data*. Statistisk sentralbyrå, Central Burau of Statistics of Norway. .

Ross, H. (2015). *Understanding and Measuring Cigarette Tax Avoidance and Evasion: A Methodological Guide*. https://www.tobacconomics.org/files/research/187/Ross_Methods_to_Measure_Illicit-Trade_03-17-15.pdf

Ross, H. (2018). Tobacco industry strategies to reduce tax liability.

Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, *2*(4), 367-374.

SPDC (Social Policy and Development Centre). (2020). Tobacco taxation in Pakistan and AJK: myth vs. reality. SPDC Policy Brief, May 2020.

Annexures

Annex A – Summary Statistics

Variables		Mean	Std. Dev.	Min	Max	Observations
lrturnover	overall	13.4829	1.6066	10.7166	15.5817	N = 112
	between		1.9251	11.3178	15.0675	n = 3
	within		0.3560	12.4597	14.3749	T-bar = 37.3333
lrprofit	overall	11.94333	2.8417	1.0848	15.3031	N = 112
	between		2.5539	9.2132	14.2981	n = 3
	within		1.9343	1.4708	13.8102	T-bar = 37.3333
lrwages	overall	10.84176	1.4540	8.4513	12.7112	N = 112
	between		1.7438	8.8761	12.2666	n = 3
	within		0.3171	10.1552	11.6937	T-bar = 37.3333
lrint_inputs	overall	12.90022	1.4659	10.3225	14.7117	N = 112
	between		1.7432	10.9147	14.2709	n = 3
	within		0.3715	12.1079	13.7164	T-bar = 37.3333

Annex B – Summary of Results Stationarity and Cointegration Tests

Stationarity Tests - Summary of the results				
Levin-Lin-Chu unit-root test Ho: Panels contain unit roots Ha: Panels are stationary Number of panels = 3 Number of periods = 36			**Im-Pesaran-Shin unit-root test** Ho: All panels contain unit roots Ha: Some panels are stationary Number of panels = 3 Number of periods = 37.33	
Variables	**Adjusted t***	**p-value**	**Z-t-tilde-bar**	**p-value**
Log Real Output	-4.4478	0.0000	-3.8881	0.0001
Log Real Wage Bill	-2.6621	0.0039	-3.9752	0.0000
Log Real Profit	-3.7421	0.0001	-5.7843	0.0000
Log Real Intermediate Inputs	-2.1239	0.0168	-2.7199	0.0033

Cointegration Tests Summary		
Variables: lrturnover lrprofit lrwages lrint_inputs		
Hypothesis		
Ho: No cointegration Ha: All panels are cointegrated		
Number of panels	=	3
	Statistic	**p-value**
Kao test for cointegration		
Modified Dickey-Fuller t	-13.111	0.000
Dickey-Fuller t	-8.028	0.000
Augmented Dickey-Fuller t	-5.890	0.000
Unadjusted modified Dickey Fuller	-15.936	0.000
Unadjusted Dickey-Fuller t	-8.184	0.000
Pedroni test for cointegration		
Modified Phillips-Perron t	-3.199	0.001
Phillips-Perron t	-9.907	0.000
Augmented Dickey-Fuller t	-7.084	0.000
Westerlund test for cointegration		
Variance ratio	-1.519	0.064
Number of periods		
Kao test for cointegration	=	35.33
Pedroni test for cointegration	=	36.33
Westerlund test for cointegration	=	37.33

Annex C – Model Estimation Results

Fixed effects model estimates

```
. xtreg $y ($xvar_prod) qdate if time <= 26, fe

Fixed-effects (within) regression               Number of obs      =        76
Group variable: company                         Number of groups   =         3

R-sq:                                           Obs per group:
     within  = 0.8117                                          min =        24
     between = 0.9996                                          avg =      25.3
     overall = 0.9890                                          max =        26

                                                F(4,69)            =     74.37
corr(u_i, Xb)  = 0.9463                         Prob > F           =    0.0000

------------------------------------------------------------------------------
  lrturnover |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   lrprofit |   .0643739   .0091281     7.05   0.000     .0461639     .082584
    lrwages |   .1504327   .1019772     1.48   0.145    -.0530062    .3538716
lrint_inputs |   .7496713   .0782233     9.58   0.000     .5936202    .9057225
      qdate |   .0117329   .0039948     2.94   0.005     .0037635    .0197024
      _cons |  -1.181586   1.869558    -0.63   0.529    -4.911252     2.54808
-------------+----------------------------------------------------------------
    sigma_u |  .15872237
    sigma_e |  .17095183
        rho |  .46295538   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0: F(2, 69) = 1.57                       Prob > F = 0.2162
```

Least-square dummy variables (LSDV) model estimates

```
. regress $y $xvar_prod pmpk ptc ktc qdate  if time <= 26
note: ktc omitted because of collinearity

      Source |       SS           df       MS      Number of obs   =        76
-------------+----------------------------------   F(6, 69)        =   1097.32
       Model |  192.412305         6  32.0687175   Prob > F        =    0.0000
    Residual |  2.01649244        69  .029224528   R-squared       =    0.9896
-------------+----------------------------------   Adj R-squared   =    0.9887
       Total |  194.428798        75  2.59238397   Root MSE        =    .17095

------------------------------------------------------------------------------
  lrturnover |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    lrprofit |   .0643739   .0091281     7.05   0.000     .0461639     .082584
     lrwages |   .1504327   .1019772     1.48   0.145    -.0530062    .3538716
lrint_inputs |   .7496713   .0782233     9.58   0.000     .5936202    .9057225
        pmpk |    .169641   .2522237     0.67   0.503    -.3335315    .6728136
         ptc |   .3171884   .3338841     0.95   0.345    -.3488922    .9832689
         ktc |          0  (omitted)
       qdate |   .0117329   .0039948     2.94   0.005     .0037635    .0197024
       _cons |  -1.348133   1.686564    -0.80   0.427    -4.712737    2.016471
------------------------------------------------------------------------------
```

Random effects model estimates

```
. xtreg $y ($xvar_prod) qdate if time <= 26, re

Random-effects GLS regression                   Number of obs      =         7
Group variable: company                         Number of groups   =

R-sq:                                           Obs per group:
     within  = 0.8102                                          min =         2
     between = 0.9997                                          avg =      25.
     overall = 0.9892                                          max =         2

                                                Wald chi2(4)       =    6477.4
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =     0.000

------------------------------------------------------------------------------
  lrturnover |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval
-------------+----------------------------------------------------------------
    lrprofit |   .0660556   .0088546     7.46   0.000     .0487009    .083410
     lrwages |   .2305432   .0728297     3.17   0.002     .0877996    .373286
lrint_inputs |   .7522091   .0729918    10.31   0.000     .6091479    .895270
       qdate |   .0141028   .0028177     5.01   0.000     .0085802    .019625
       _cons |  -2.631789   .7172339    -3.67   0.000    -4.037541   -1.22603
-------------+----------------------------------------------------------------
     sigma_u |          0
     sigma_e |  .17095183
         rho |          0   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Ordinary least squares (OLS) pooled regression model estimates

```
. regress $y $xvar_prod qdate  if time <= 26

      Source |       SS           df       MS      Number of obs   =         7
-------------+----------------------------------   F(4, 71)        =    1619.3
       Model |  192.32076          4  48.0801899   Prob > F        =    0.000
    Residual |  2.10803816        71  .029690678   R-squared       =    0.989
-------------+----------------------------------   Adj R-squared   =    0.988
       Total |  194.428798        75  2.59238397   Root MSE        =    .1723

------------------------------------------------------------------------------
  lrturnover |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval
-------------+----------------------------------------------------------------
    lrprofit |   .0660556   .0088546     7.46   0.000     .0484001    .083711
     lrwages |   .2305432   .0728297     3.17   0.002     .0853249    .375761
lrint_inputs |   .7522091   .0729918    10.31   0.000     .6066677    .897750
       qdate |   .0141028   .0028177     5.01   0.000     .0084845    .019721
       _cons |  -2.631789   .7172339    -3.67   0.000    -4.061912   -1.20166
------------------------------------------------------------------------------
```

Annex D – Robustness Model Estimation Results

Quarterly Dummies Model estimates

regress $y $xvar_prod qdate q1 q3 q4 if time <= 26						
					Number of obs =	76
Source	SS	df	MS		F(7,68) =	1284.44
Model	192.29	7	27.47		Prob > F =	0.0000
Residual	1.45	68	0.02		R-squared =	0.9925
Total	**193.74**	**75**	**2.58**		Adj R-squared =	0.9917
					Root MSE =	0.1462
lrturnover	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
lrprofit	0.0563	0.0078	7.20	0.000	0.0407	0.0719
lrwages	0.2637	0.0767	3.44	0.001	0.1107	0.4167
lrint_inputs	0.7301	0.0768	9.51	0.000	0.5769	0.8834
qdate	0.0135	0.0024	5.60	0.000	0.0087	0.0183
q1	-0.1879	0.0485	-3.88	0.000	-0.2846	-0.0912
q3	-0.1529	0.0494	-3.09	0.003	-0.2515	-0.0543
q4	-0.2008	0.0539	-3.73	0.000	-0.3083	-0.0933
_cons	-2.3192	0.6287	-3.69	0.000	-3.5738	-1.0647

Outlier Dummy model estimates

regress $y $xvar_prod qdate out_dum if (time <= 26)						
					Number of obs =	76
Source	SS	df	MS		F(5,70) =	2419.24
Model	192.62	5	38.52		Prob > F =	0.0000
Residual	1.11	70	0.02		R-squared =	0.9942
Total	**193.74**	**75**	**2.58**		Adj R-squared =	0.9938
					Root MSE =	0.1262
lrturnover	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
lrprofit	0.0566	0.0066	8.59	0.000	0.0435	0.0698
lrwages	0.1506	0.0543	2.77	0.007	0.0423	0.2589
lrint_inputs	0.8454	0.0548	15.43	0.000	0.7361	0.9547
qdate	0.0132	0.0021	6.38	0.000	0.0091	0.0173
out_dum	0.5510	0.0775	7.11	0.000	0.3964	0.7057
_cons	-2.6760	0.5253	-5.09	0.000	-3.7237	-1.6283

Random effects model estimates

regress $y $xvar_prod qdate q1 q3 q4 out_dum if time <= 26						
					Number of obs =	76
Source	SS	df	MS		F(8,67) =	1722.01
Model	192.80	8	24.10		Prob > F =	0.0000
Residual	0.94	67	0.01		R-squared =	0.9952
Total	**193.74**	**75**	**2.58**		Adj R-squared =	0.9946
					Root MSE =	0.1183
lrturnover	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
lrprofit	0.0524	0.0064	8.25	0.000	0.0398	0.0651
lrwages	0.1507	0.0648	2.33	0.023	0.0214	0.2799
lrint_inputs	0.8505	0.0652	13.04	0.000	0.7203	0.9807
qdate	0.0132	0.0019	6.77	0.000	0.0093	0.0171
out_dum	0.4758	0.0783	6.08	0.000	0.3195	0.6321
q1	-0.1385	0.0400	-3.46	0.001	-0.2184	-0.0586
q3	-0.0894	0.0413	-2.16	0.034	-0.1719	-0.0069
q4	-0.0985	0.0467	-2.11	0.039	-0.1918	-0.0053
_cons	-2.6086	0.5108	-5.11	0.000	-3.6282	-1.5891