# PEDL Research Papers

This research was partly or entirely supported by funding from the research initiative Private Enterprise Development in Low-Income Countries (PEDL), a Foreign, Commonwealth & Development Office (FCDO) funded programme run by the Centre for Economic Policy Research (CEPR).

**Productivity and Reallocation under Monopolistic Competition: A Micro Panel Data Analysis**

**Samuel V. Mhlanga**[1]

**Final Version: 5 August 2023**

**Abstract**

This article studies the structural aggregate productivity growth (APG) decomposition with demand- and supply-side controls, determines comparative statics predictions for firms and economic outcomes, and examines patterns of input distortions. By moving from price-taking conditions to markets featuring markup heterogeneity for product varieties, the paper finds amplification of production inefficiency from -3.61% to -11.41% and amplification of total factor reallocation from 0.15% to 8.91%. The productivity results are robust to structural variations in the demand function, firm scale adjustment, and firm growth. Similarly, input reallocation is robust to variation in demand structure and plant expansion. Furthermore, reallocation under *common markups* among all firms is robust to reallocation under *heterogeneous markups* among larger firms. Alternatively, large firms face demand inelasticities and charge higher markups thereby mimicking the behaviour of the survey of all firms. Under autarky, small unproductive plants charge higher markups than their small efficient counterparts. Demand elasticity increases (decreases) with industry output for smaller (larger) plants. Finally, a unit-increase in capital-intensity for resource-constrained plants raises labour distortions and reduces capital distortions while reducing capital distortions for resource-unconstrained firms.

*JEL*: D22, D24, L6

*Keywords*: Productivity, Reallocation, Monopolistic Competition, Heterogeneity, Demand Curvature, Input Distortions, Eswatini

[1] I acknowledge insightful comments from Chris Fowler and colleagues at the University of Eswatini, my discussant Niclas Moneke and conference participants at CSAE, University of Oxford, UK. I also thank an anonymous referee and the associate editor of this journal for suggestions. I acknowledge with thanks the Exploratory Research Grant (Ref. 5765) provided by the Private Enterprise Development in Low-Income Countries (PEDL); a research initiative of the Centre for Economic Policy Research (CEPR) and the Department of International Development (DFID), UK. I am indebted to the Central Statistical Office of Eswatini for providing firm-level data to carry out this work. The standard disclaimer remains.

## 1. Introduction

A central focus in industrial organization is the differential impact between productivity and cross-firm input reallocation on aggregate productivity growth (APG).[2] A widely cherished insight featuring productivity dispersion is the Darwinian process of inefficient plants shrinking and exiting the industry while efficient ones survive and grow (Petrin et al., 2011; Nishida et al., 2014; Kwon et al., 2015). Moreover, firm-level differences in supply and demand characteristics are also notable drivers of variation in productivity growth. To capture the effects of reallocation and productivity on APG, the literature has thus far relied almost exclusively on price-taking assumptions. An associated robust finding associated with zero pricing power is that input reallocation among incumbents and entry/exit dynamics shape variations in industry aggregates. This factor-input churning has attracted theoretical and empirical attention since the early 1980s. However, the neglect of industrial product differentiation and markup variability deprives related agents the opportunity to understand the real impact of pricing power on productivity and reallocation under imperfect competition, a primary theme of this paper.

The purpose of this article is to decompose the structural APG under price-taking and price-setting conditions. In the process; it determines the sensitivity of productivity and input reallocation to variations in firm scale and changes in the structure of demand faced by plants, while price-taking outcomes serve as baseline results for comparison with the mainstream evidence. To the best of our knowledge, this is the first paper to measure and examine plant-specific markups to transform the APG decomposition price-taking results to markup pricing outcomes. In exploring flexible consumer preferences and their effects on changes in the APG decomposition, some comparative statics predictions emerge between demand parameters and both firm-behaviour and economic performance. At the same time, the measurement of markups and responsiveness of output quantity demanded to changes in prices (or demand elasticity) is undertaken for various categories of plants. The paper ends with briefly relating producer characteristics to changes in capital and labour distortions that contaminate the reallocation component of APG.

The bedrock of our analysis is that the selection mechanism in firm dynamics characterizes industries as groups of firms with varied productivity levels, and relates this productivity to performance and survival of plants in the industry.[3] The pioneering contributions in this effort are Jovanovic (1982), Hopenhayn (1992), Ericson and Pakes (1995), Melitz (2003), and Asplumd and Nocke (2006).[4] A critical channel for APG in

[2] The words firm, plant, producer, and establishment are used interchangeably due to the data structure at hand.

[3] Productivity is understood as referring to a firm's act of hiring inputs for the production of output. High productivity or production efficiency firms are those that hire low factor-inputs for the production of higher output, sometimes also called low-cost firms.

[4] More recently, Weintraub et al. (2011) propose an approximation method for analysing finite dynamic models of imperfect competition of the Ericson and Pakes (1995) type, propose foundations for infinite models of the Hopenhayn (1992) type while providing an alternative equilibrium concept and computational methods for the new solution concept.

these models is the market-share shift across incumbent plants and entry/exit margins. It is now commonly accepted that the propensity for low productivity plants to fail relative to their more productive counterparts induces selection-driven adjustment in APG. Therefore, the productivity-survival connection is theoretically a critical force giving full thrust to the whole APG decomposition concept.

Empirically, the productivity-survival nexus has been confirmed. For instance, Petrin et al. (2011), Nishida et al. (2014), and Kwon et al. (2015) found that producers with poor supply fundamentals under price-taking conditions have higher propensity to exit and reallocate their market shares to stronger incumbents and entrants. However, as argued by Syverson (2011), Foster et al. (2016) and Foster et al. (2017), both supply- and demand-side shocks matter for firm-level decisions to invest or shut down operations. Ultimately, the aggregate effect of unconstrained firm churning is that input reallocation dominates technical progress, with a few exceptions (e.g., Kwon et al., 2015).

Furthermore, there are two important parameters for understanding comparative statics predictions in the preferred price-setting environment: the demand elasticity and curvature of the inverse demand function. This framework introduces consumer behaviour that sets prices to marginal utility (Foster et al., 2018; Dhingra and Morrow, 2019; and Matsuyama and Ushchev, 2022). With the standard incomplete pass-through of marginal costs to prices, an increase in the elasticity of inverse demand is positively correlated with prices and therefore markups of product varieties. Under well-behaved consumer preferences, markups are also an increasing function of output and therefore production technologies embodied in output. The demand elasticity also tends to be large for small plants, but declines with an increase in sales (Burya and Mishra, 2022). Equally importantly, scaling productivity and revenue function elasticities with the parameter that defines the demand elasticity (hereafter referred to as the demand parameter) amplifies baseline performance measures in varying degrees (Foster et al., 2018).

This paper outlines the theoretical underpinnings of the structural accounting decomposition of APG and shows how to quantify it using micro panel data from the manufacturing sector at Eswatini. Our preference for this framework lies in its theoretical strengths. First, it rationalises the contributions of micro productivity growth, resource reallocation across plants, and firm entry/exit dynamics on APG using sound microfoundations. Second, the decomposition depends crucially on consistent estimation of revenue function coefficients and back out unbiased regression residuals. In particular, any reallocation term is a function of input growth, and the difference between the value of marginal products of factor-inputs (VMPs) and the marginal costs of these inputs (Petrin and Levinsohn, 2012; and Petrin et al., 2011). Although the *raison d'être* for developing the structural APG decomposition framework was to estimate physical output elasticities, proxy methods are however designed to estimate revenue function elasticities under price-taking conditions. The residual from the revenue production technology is a function of fundamentals of output-based Total Factor Productivity ($TFPQ$) and its growth as well as demand shifters

(Foster et al., 2017). This residual and related revenue function elasticities require scaling by a demand parameter to recover an APG that is based on technical efficiency and physical output elasticities.

Our empirical approach to estimating revenue function elasticities and regression residuals derives from control function methods applied to production technologies while avoiding the simultaneity biases of Ackerberg et al. (2015). In particular, if observed factor-inputs are a function of unobserved determinants of output as in the first stages considered in Olley and Pakes (1996) and Levinsohn and Petrin (2003), then there is an endogeneity problem in the estimation procedure. Wooldridge (2009) modifies the Levinsohn-Petrin Estimator to produce consistent results within a single-step in a Generalized Method of Moments (GMM) framework. In this study, we follow Petrin et al. (2011) and Nishida et al. (2014) who adopt the Wooldridge-Levinsohn-Petrin Estimator to avoid the Ackerberg et al. (2015) criticism.

This paper shares an obvious common thread with Petrin and Levinsohn (2012) and Petrin et al. (2011). Its applied dimension is closest to Nishida et al. (2014) and Petrin et al. (2011) for price-taking baseline results; Foster et al. (2017) and Foster et al. (2018) for estimation under price-setting conditions, and its demand structure derives from Dhingra and Morrow (2019). It departs from the empirical literature by extending Foster et al. (2017), recasting the decomposition under monopolistic competition with firm-level heterogeneity in markups. We introduce the demand side by assuming firms face Hyperbolic Absolute Risk Aversion (HARA) preferences, and Constant Elasticity of Substitution (CES) utility technologies as a special case, due to its tractability and because such preferences are an essential restriction of economic analysis (Perets and Yashiv, 2015). This allows for systematic variation in the decomposition after a scaling procedure recovering physical-quantity elasticities and technical efficiency. Linked to the HARA utility choice is that productivity dispersion provides a mechanism for efficient input reallocation and complements the notion that rich demand systems such as the HARA utility support deeper learning about market outcomes (Mrázová and Neary, 2017).

A preview of the results under price-taking conditions shows deterioration in production efficiency and miniscule input reallocation gains from incumbent firms. Under monopolistic competition; however, technical efficiency remains significantly attenuated and factor-input reallocation amplified. Technical efficiency is robust to further variations in the architecture of the demand function. Although the pace of factor-input reallocation is also robust to demand variations, it breaks down with respect to firm-size due to reduced paid employee reallocation for smaller plants and reduced capital reallocation for larger plants. Furthermore, there is generally inelastic demand for product varieties while corresponding markups are higher. Moreover, a unit-increase in capital-intensity for resource-constrained plants in manufacturing raises labour distortions and reduces capital distortions. That same increase in capital-intensity; *albeit* for resource-unconstrained firms, reduces capital distortions.

This paper makes at least three contributions to the literature. First, it produces new results on the APG decomposition for an under-explored Sub-Saharan small open economy. Second, it introduces granular assumptions of monopolistic competition to the structural accounting decomposition of APG. Third, it determines product elasticities and markups by firm size and growth category for productive and unproductive plants.

The organization of this paper is as follows. The next section presents an overview of the manufacturing sector in Eswatini, discussing the state of product complexity versus structural transformation. Section 3 summarizes the theoretical setup of the APG structural decomposition. Section 4 reviews the micro dataset and establishes its quality. Section 5 reports APG decomposition results under price-taking conditions, introduces the monopolistic competition framework, presents comparative statics predictions, and reports APG decomposition results under monopolistic competition and factor-input distortion dynamics. Section 6 summarises and concludes.

## 2. Overview of the Manufacturing Sector in Eswatini

To understand the state of manufacturing in Eswatini, it seems appropriate to outline a framework for discussing the sector's technological content, economic growth and development. A good place to start concerns the evolution of gross domestic product (GDP) as explained by the structure of labour and capital markets as well as their individual productivity, collectively referred to as productive capabilities. Because economic growth and structural transformation are intrinsically connected through the architecture of national productive capabilities, we rely on a stylized specification of production technology for annual GDP generation.[5] In this context, structural transformation denotes the reallocation of productive resources from low- to high-economic activities across broadly defined sectors consisting Agriculture, Manufacturing, and Services (Herrendorf et al., 2013). The evidence sourced for other economies is that sectoral heterogeneity in technical change is the main factor driving variation in structural transformation (Herrendorf et al., 2015). Given constant variations in economic technical change, institutional quality and productive capabilities embodied in factor-inputs; production technologies are deployed to generate GDP that evolves according to

$$GDP_t = Q_t(A_K K_t, A_L L_t) exp(\omega_t + \vartheta_t), \forall t \in [1994, 2003], \quad [1]$$

where $\omega_t$ refers to a Hicks-neutral productivity shock, $A_K$ is capital-augmenting technical change, and $A_L$ is labour-augmenting technical change. Eq. 1 presents the stock of capital as $K_t$ and the stock of labour is $L_t$, while $\vartheta_t$ denotes random noise. Thus, long-run and sustainable $\Delta lnGDP_t$ (or economic growth) is an outcome of continuous technological improvement, and product selection for production while trade is an outcome of revealed comparative advantage (RCA) in the arguments of Eq.1. Consequently, the annual

[5] The terms *structural transformation* and *structural change* are used interchangeably to refer to the movement of market shares of resources across economic sectors.

rate of economic growth for Eswatini arising from factor-input mix was on average 3.3% during the *de facto* trade liberalization period of 1994-2003 in the Southern African Customs Union (SACU).

A robust stylized fact in the organization of production is the secular decline in the share of agricultural output in GDP versus growth in output shares of the other broad sectors. The economy of Eswatini is by no means different. The average size of the Manufacturing sector was 35.3% relative to national GDP and declining in the 1995-2005 period. In terms of the number of industries, the sector had 13 Two-Digit ISIC broad industries and 49 Four-Digit ISIC narrowly defined industries, already experiencing premature deindustrialization in the reference period (UNECA, 2018). As highlighted by Edwards and Behar (2006), the intensity of import competition induced by *de facto* trade reforms facilitated the movement of market shares of factor-inputs to higher-productivity firms while creating a suitable environment for foreign technology diffusion to domestic producers. As a result, Gross Value Added (GVA) in Manufacturing rose by three percent, with the Services sector increasing its GVA by seven percent while Agriculture declined by circa nine percent.[6]

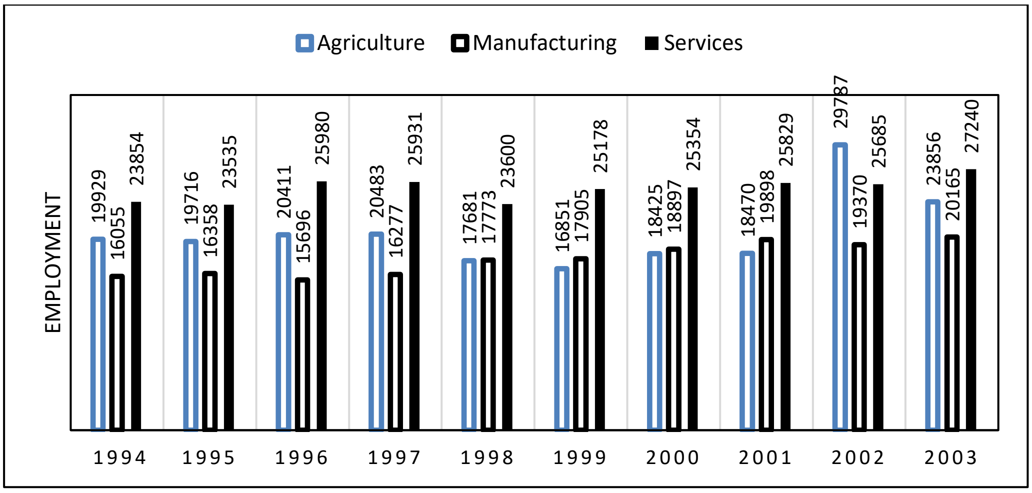



**Figure 1: Distribution of Sectoral Employment from 1994-2003**

**Source: UNECA (2018)**

Furthermore, employment trends in the sector as shown in Figure 1 mimic patterns observed in Africa and in other parts of the world. As De Vries et al. (2015) and McMillan and Zeufack (2022) deduce in their study of African countries, the continent's employment record takes a robustly increasing trend only post-2000. Similarly, the labour market in Eswatini remained somewhat subdued in the first three years and only experienced marginal expansion from 1998. To strengthen and expand its base as well as its transition into high-technology products, the sector requires access to capital investment in a form of either FDI or local provision or both. In its post-2000 period, gross capital formation declined from 19% to 12%. These capital

[6] See UNECA (2018).

inflow patterns coincided with a decline in consumption, which jointly accounted for *part* of the fall in economic growth (UNECA, 2018).

In the same period, the export basket for the sector was dominated by the same basic exports of pre-1968 National Independence; i.e., exports of sugar, forest products, meat products, and citrus fruits. In the passage of time, the export mix evolved to include an additional cluster of standard products; namely, apparel products, processed fruits, and soft drinks by 2003. As noted by Hartmann et al. (2017), a country's product mix predicts its pattern of consequential product diversification and $\Delta lnGDP_t$. Thus, $\Delta lnGDP_t$ is largely an outcome of product upgrading and export diversification. The national capabilities necessary for making new products are readily adaptable for some products than for others. UNECA (2018) found that export sales declined significantly from 1999 to 2001, as also shown in Figure 2, and by 18% in 2000-2009. At the same time, medium- to high-technology products accounted for only 14% of total exports in 1990 and 29% by 2007. Clearly, this level of technological content in the product group demonstrates that the existing complexity of the export basket is neither a result of deep structural transformation nor that of technologically-intensive processes.

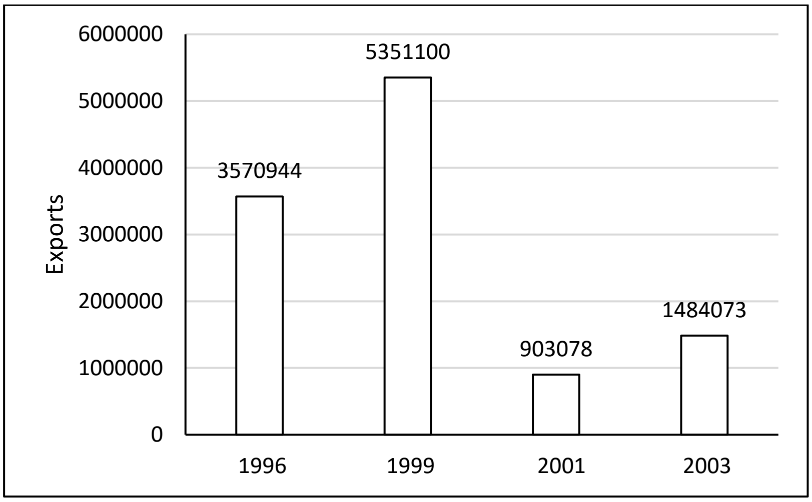



**Figure 2: Distribution of Exports Covering the Period 1996-2003 in Eswatini (Emalangeni)**

**Source**: IMF Country Reports

Product diversification in pursuit of proximity of the national export basket to the global product space and realignment of associated resource capabilities are important for $\Delta lnGDP_t$. Production and exportation of sophisticated goods need to move to the global export frontier so that the export concentration ratio and export diversification index collectively decline. For instance, the Export Concentration Ratio $(EC_R)$ has been 0.25 while sugar accounted for 11% of total exports since 1995 in Eswatini, compared to at least

0.5% for most African economies.[7] The $EC_R$ measures the degree to which a large share of exports is represented by a few products. On the other hand, a high Export Diversification Index $(ED_I)$ of 0.75 in the same period indicates a greater divergence of the distribution of Eswatini's exports from global patterns (UNECA, 2018).[8] The weak complexity of the export mix resonates with the poor economic fundamentals reported in UNECA (2018). That is, Eswatini experienced the weakening of capital investment, stock of skilled labour as well as in institutional quality. Consequently, 'within-sector' productivity growth weighted by employment share reckoned from the beginning of the period accounted for 0.32% in Figure 3 compared to 0.50% in South Africa in the period 1990-2000 (de Vries et al., 2015). The 0.58% structural transformation partly reflects constrained productive capabilities and other economic fundamentals that shift an economy to more 'complex' export products. Such bottlenecks can be undone through continuous product adjustment and resource adaptation in response to product diversification.


**Figure 3: Structural Transformation and Within-Sector Productivity Growth for World Economies**

**Source: McMillan, Rodrik, and Verduzco-Gallo (2014) and UNECA (2018).**

---

[7] This Export Concentration Ratio is calculated as, $EC_R \in [0,1] = \left(\sqrt{\sum_l^N \left(\frac{X_{i,g}}{X_g}\right)^2} - \sqrt{\frac{1}{N}}\right) \times \left(1 - \sqrt{\frac{1}{N}}\right)^{-1}$, where $g$ is the exporting country, $X_{l,g}$ is the value of export product $l$ from country $g$, and $N$ denotes the product range of exports at the three-digit SITC level Revision 3. When $EC_R = 1$, it means all exports from country $g$ are accounted for by a single commodity. On the other hand, when $EC_R = 0$, it means the country's exports are evenly distributed across all its products, Cadot, Carreŕe, and Strauss-Kahn (2011).

[8] The computation of the Export Diversification Index is $ED_I \in [0,1] = \frac{1}{2}\sum_l^N \left|h_{lg} - x_i\right|$, where $h_{lg}$ is the share of commodity $l$ in the total exports of country $g$ and $x_l$ is the share of the commodity in world exports. Notably, $ED_I$ increases with a fall in export diversification since the higher the index the more disperse the deviation is of the country's exports from the global export patterns. Cadot, Carreŕe, and Strauss-Kahn (2011) introduce the normalized Herfindahl, Gini, and Theil's entropy indices to measure export concentration and diversification.

Looking at continental and regional productivity outcomes, the results from Africa, Latin America and High-Income countries (except for Asia) show negative patterns of structural change, indicating adverse movement of resources. Patterns of historical variation in structural transformation concerning economies and sectors are common in Africa and elsewhere, and cross-sectional variations in countries' structural change are also prevalent (de Vries et al., 2015; de Vries et al., 2016; and McMillan et al., 2014). Not surprisingly, both within-sector and structural change in Figure 3 for Eswatini are weakly positive, reflecting some economy-wide prospect of sustained economic growth and modest movement of productive resources in the 'right' direction.

Therefore, the motivation for this piece is the cross-movement of resources from low- to high-activity uses that consequently boost $\Delta lnGDP_t$. As such, the stronger the structural transformation is in the economy, the stronger the implied complexity of the export product mix. With efficient structural change, the closer the national products to the core of the global product space, the higher the income generation prowess for participating economies (Hausmann et al., 2007). For the case of Eswatini, is structural change moving in the right direction, or otherwise, given the 'almost' *static* complexity of export products and *static* RCA? Does monopolistic competition vary the pace of structural change, regardless of the direction of resource flows? What are firm-specific characteristics that influence changes in market distortions for factor-inputs? Indeed, de Vries et al. (2015) insist on; *inter alia*, deeper analyses of the impact of technological progress and factor reallocation on APG. This is carried out in the next sections.

## 3. The Setup

This section focuses on the structural accounting framework for the contribution of plant-level technological progress, input reallocation, and entry/exit margins to APG. To this end, it lays out the theoretical foundations developed by Petrin and Levinsohn (2012)/Petrin et al. (2011) (hereafter referred to as PL/PWR) who define APG as the change in aggregate final demand *minus* the change in aggregate factor input expenditures. While the underlying PL/PWR mechanism allows for aggregating idiosyncratic measured quantities to aggregate outcomes, it also allows APG to increase without input adjustments, conditional on plants becoming more technically efficient.

The next two subsections therefore concentrate on the preferred accounting decomposition for APG based on consistently estimated production function parameters for each producer. We then show how to aggregate microeconomic dynamics of incumbent firms, entrants, and exiting plants to macroeconomic APG outcomes.

### 3.1. Continuous-Time Aggregation of Productivity Growth

Our preferred approach is the structural accounting decomposition pioneered by PL/PWR due to its sound microfoundations. This section dissects the structural APG composition to learn about the direction of change in response to changes in within-plant productivity, input reallocation, and plant turnover. The approach is attractive for several reasons. First, the method delivers a mechanism for contributions arising from firm-level productivity, factor-input reallocation among incumbents, and entry/exit margins to APG. Second, the framework also confronts characteristics that drive variation in plant-level data such as idiosyncratic pricing power. Third, the decomposition is based on scientific estimation of production technologies to extract consistent revenue function and/or physical output elasticities. A fundamental and attractive property of the structural APG decomposition lies in its independence from the choice of productivity estimator or assumptions about market structure, given the definition of APG (Foster et al., 2017). Therefore, comparisons of within-plant effects and factor-input reallocations across various forms of pricing behaviour exhibited by producers therefore remain feasible.

As PL and PWR make clear, perhaps a key insight arises from the exploration of reallocation implications of inputs across incumbent plants. As it turns out, factor-input reallocation across firms is a function of factor-input growth and gaps between the values of marginal product of inputs (VMPs) and marginal costs (MCs), hereafter referred to as the (VMP-MC) gap. The measured gap itself depends critically on the extent of measurement precision and consistent estimation of output elasticities.

Thus, consider the plant-specific production technology similar to Eq. 1 that gives rise to firm-level output for use in the APG structural decomposition

$$Q_{it} = Q_{it}(K_{it}, L_{it}, M_{it}, \Omega_{it}), \forall i \in N \text{ and } \forall t \in T, \qquad [2]$$

where $Q_{it}$ denotes the physical quantity of output and the arguments therein chronologically refer to the stock of capital $K_{it}$, the stock of labour $L_{it}$, the stock of intermediate inputs $M_{it}$, and the Hicks-neutral log-additive idiosyncratic multifactor productivity shock, $\Omega_{it}$, for plant $i$ at time $t$.[9]

Rather than referring to multifactor productivity as a measure of our ignorance, the literature on productivity and structural transformation has very specific characterizations of total factor productivity (Syverson, 2011; Hausmann, 2016). The former distinguishes between internal and external levers to the firm, while Hausmann's definition of the concept identifies clusters of productivity-enhancing capabilities

[9] Differences in technical efficiency can be factor-augmenting in some settings. For example, the characterization of production technology, $[Y_i]$, may involve a common CES between factor-inputs, $\sigma$, for a set of N plants. For example, firm $i$'s production function can be $Y_i = A\left[(A_K K_i)^{\frac{\sigma-1}{\sigma}} + (A_L L_i)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$, where **A** is Hicks-neutral log-additive technical efficiency unobservable to the econometrician but observable to the firm, $A_K$ is plant $i$'s capital stock augmenting technical efficiency and $A_L$ is labour-augmenting technical change. When $\sigma = 1, \sigma = 0,$ or $\sigma = \infty$, then $Y_i$ takes the Cobb-Douglas, Leontif, or linear functional form, respectively; although we do not discuss these individually for ease of readability. Alternatively; $Y_i$ can also take factor-augmenting Constant Elasticity of Substitution (CES)-nested functional form as in Oberfield, and Raval (2021), Raval (2019), and Demirer (2022).

deployed by firms.[10] In general, this implies that each firm in an industry can have different sets of production technologies built around a combination of these levers or clusters. Related industries, in turn, are presumably populated by single-product firms. A property like this follows because producers do not allocate different factor-inputs to the production of different product varieties.

In the structural decomposition framework; therefore, the output amount that goes into aggregate final demand is $Y_{it}$. Then $Y_{it} = Q_{it} - \sum_j^J M_{ijt}$, where $M_{ijt}$ are intermediate input expenditures for firm $i$ consumed by firm $i$ or firm $j$ or both at time $t$, depending on whether or not firm $i$ is vertically integrated. In the make or buy decision concerning intermediates; for instance, firm $i$ may choose to make its own material inputs thereby achieving vertical integration while at the same time avoiding inefficiencies of suppliers of intermediate inputs. Thus, after some manipulation of variables and reduction of clutter by dropping the time subscript, we assume $Q_{it}$ is differentiable in order to derive APG initially in continuous-time as

$$APG = \sum_i \sum_m \left( P_i \frac{\partial Q_i}{\partial X_m} - W_{im} \right) dX_{im} + \sum_i \sum_j \left( P_i \frac{\partial Q_i}{\partial M_j} - P_j \right) dM_{ij} + \sum_i P_i \frac{\partial Q_i}{\partial \Omega_i} d\Omega_i + \sum_i dNE_i \qquad [3]$$

where $m \in (L, K)$ denotes factor-inputs that define stocks of capital $X_{iK}$ and labour $X_{iL}$ for firm $i$. The quantities $P_i \frac{\partial Q_i}{\partial X_m}$ and $P_i \frac{\partial Q_i}{\partial M_j}$ are the respective VMPs of factor-inputs and intermediates with $P_i$ as the unit price of output $Q_i$ for firm $i$, while $W_{im}$ and $P_j$ are the respective factor and intermediate input prices that reflect marginal costs. The first two terms in Eq. 3 reflect aggregate productivity contributions by incumbent plants through factor and material input reallocation, respectively. The third term is a measure of aggregate technical efficiency while the final term represents the net-entry contribution to APG. This model clearly abstracts from fixed and/or sunk costs in order to focus attention on reallocation, technical efficiency and the entry/exit dynamic of plants.

Eq.3 directly reflects the fundamental structure of the accounting decomposition of APG. It shows that if the (VMP-MC) gaps are equalized, then two polar outcomes are expected. First, in the presence of very high input market frictions or adjustment costs, input reallocation may completely be hindered. Second, gap equalization across all firms and inputs allows for *a priori* continuous cross-firm reallocation in response to infinitesimal perturbations to the economic system in order to maintain the effects of prevailing market

[10] The internal 'levers' in Syverson (2011) include managerial talent, high quality factor-inputs, information technology and R&D, learning by doing, product innovation, and firm structure decisions. Syverson (2011) identifies complementary external sources of 'levers' of multifactor productivity as productivity spill-overs, intermarket and trade competition, regulation or proper regulation. On the other hand, the literature on structural transformation posits three classes of multifactor productivity; namely: 1) tools, embodied knowledge and/or recipes; 2) blueprints and/or codified knowledge, and 3) knowhow and/or tacit knowledge (Hausmann, 2016). Hausmann (2016) uses metaphors to demonstrate that, although the first two sources of technical progress are readily sharable within a community or a society, technical knowhow is more difficult to impart to the next user. One way to foster accumulation of tacit knowledge/knowhow over time is by repeated imitation of existing productive technologies through learning by doing/watching.

conditions. A positive gap, for instance, means factor-inputs and/or intermediates move from low- to high-production units thereby raising APG without an increase in factor-inputs. As De Loecker et al. (2020) note, heterogeneity in price-cost margins may induce movement of inputs from low- to high-markup plants.

Furthermore, PL work out a collorary to allow for recasting Eq. 3 in terms of growth rates in a continuous-time, value-added Cobb-Douglas framework as follows

$$APG = \left( \overbrace{\sum_i \sum_m D_i^v \left(\alpha_{im}^v - s_{im}^v\right) d\, ln\, X_{im}}^{Input\, Re\, allocation} \right) + \left( \overbrace{\sum_i \sum_j D_i^v \left(\alpha_{ij}^v - s_{ij}^v\right) d\, ln\, M_{ij}}^{Material\, Re\, allocation} \right) + \left( \overbrace{\sum_i D_{it}^v d\, ln\, \Omega_i}^{Technical\, Efficiency} \right) + \left( \overbrace{\sum_i dNE_i}^{NetEntry} \right) \quad [4]$$

where $D_i^v \coloneqq \frac{P_i Q_i}{\sum_{i=1}^N VA_i}$, also known as the Domar-weight arising from the Hulten's Theorem, physical-output elasticities with respect to factor and intermediate inputs are $\alpha_{im}^v$ and $\alpha_{ij}^v$, and total cost-shares for input expenditures are $s_{im}^v = \frac{W_{im} X_{im}}{VA_i}$ and $s_{ij}^v = \frac{P_j M_{ij}}{VA_i}$.[11]

As with the (VMP-MC) gap, the equivalent $(\alpha_{im}^v - s_{im}^v)$ gap is still the basis for reshuffling inputs across establishments. For any factor-input $X_{im} \in (X_{iK}, X_{iL})$, the related reallocation term is $\sum_i \sum_m D_i^v (\alpha_{im}^v - s_{im}^v) d\, ln\, X_{im}$ and is at least positive if $\alpha_{im}^v \geq s_{im}^v$ and $d\, ln\, X_{im} \geq 0$ for expanding plants or $\alpha_{im}^v \leq s_{im}^v$ and $d\, ln\, X_{im} \leq 0$ for downsizing establishments. The material reallocation term, $\sum_i \sum_j D_i^v \left(\alpha_{ij}^v - s_{ij}^v\right) d\, ln\, M_{ij}$, bears the same interpretation, with $M_{ij}$ as defined for equation 3. The last term, $\sum_i dNE_i$, is net-entry. This term reflects firm growth as a difference between firms that do not exist at time $t-1$ but exist at time $t$ and those that exist at time $t-1$ but vanish at time $t$.

The technical efficiency term, $\sum_i D_{it}^v d\, ln\, \Omega_i$, represents $lnTFPQ$ because input elasticities refer to physical quantities of the production technology. This reflects aggregate growth in the multifactor productivity discussed hitherto. As will become clear, the characterization of within-firm effects according to a type of multifactor productivity (also known as total factor productivity, or *TFP*) depends critically on the estimation of output elasticities. That is, the distinction here is between revenue function and physical output-based *TFP*. More specifically, this productivity depends on whether factor input elasticities are based

[11] The generality of the Hulten's theorem derives from the envelop conditions of the first welfare theorem under *perfect competition* and *distortion-free* market environments. The said Theorem is used by Hulten (1978) and Basu and Ferdinand (2002) to equate aggregate value-added with aggregate final demand since microeconomic intermediate input usage cancels out at the aggregate level. This gels well with the irrelevance condition by Lucas (1977) that microeconomic shocks have no macroeconomic effects because they cancel out at the aggregate. However, this is in contrast with Gabaix (2011) who works with the granularity hypothesis to discount the Lucas argument and, by extension, the idea that microeconomic details do not matter for macroeconomic outcomes.

on physical output or revenue production functions. In the long run, input cost-shares in Eq.4 become immutable cross-plant average estimates of physical-output elasticities, conditional on cost-minimization and constant returns to scale (CRS) assumptions governing production technologies (Foster et al., 2017).

Prior to taking the APG decomposition to the data; however, it is necessary to move from continuous- to discrete-time aggregation of productivity growth. We do this in the next section.

### 3.2 Discrete-Time Aggregation of Productivity Growth

Given the discrete nature of micro-panel datasets, the expression is readily estimable using the Törnqvist-Divisia approximation where prices contained in Domar-weights are updated through a process of annual chain-weighting as in Nishida et al. (2014). The discrete form of the expression is ultimately re-stated as

$$APG \cong \left(\overbrace{\textstyle\sum_i \overline{D}_i^v \sum_k \left(\beta_{im}^v - \overline{s}_{im}^v\right) \Delta \ln X_{im}}^{Input\ Re\ allocation}\right) + \left(\overbrace{\textstyle\sum_i \overline{D}_i^v \sum_j D_i^v \left(\beta_{ij}^v - s_{ij}^v\right) \Delta \ln M_{ij}}^{Material\ Re\ allocation}\right) +$$

$$\left(\overbrace{\textstyle\sum_i \overline{D}_i^v \Delta \ln \Omega_i}^{Total\ Factor\ Productivity}\right) + \left(\overbrace{\textstyle\sum_{i \in Ent} D_{it}\left[1 - \sum_m s_{im}\right] - \sum_{i \in \chi} D_{it-1}\left[1 - \sum_m s_{im}\right]}^{NetEntry}\right) \qquad [5]$$

where $\overline{D}_i^v = \frac{D_i + D_{i-1}}{2}$, and $\overline{s}_{im}^v = \frac{s_{im}^v + s_{i-1}^v}{2}$.

Eq.5 presents a version of the APG decomposition that is directly estimable under price-taking and price-setting market assumptions, which we refer to as the Estimating Equation. In particular, the full application of the Estimating Equation is found in Section 5. To explain its component parts, the equation has $\beta_{im}^v$ and $\beta_{ij}^v$ because the related production technology is estimated using a Hicks-neutral Cobb-Douglass revenue function. The symbols represent revenue-function elasticities. The difference operator $\Delta$ means $\Delta \ln X_{imt} = \ln X_{imt} - \ln X_{imt-1}$, and $Ent$ denotes entrants whereas $\chi$ refers to exiters.

To capture the intuition behind the Estimating Equation under price-taking conditions, *at least* two polar cases emerge. Consider first a firm with labour elasticity $\beta_{iLt}^v = 0.45$, labour cost-share of value-added $s_{iL}^v = 0.012$ and its $t-1$ lag value of $s_{iL-1}^v = 0.009$ as performance characteristics, where the subscript $L$ represents labour. At a growth rate of $ln(110) - ln(95) = 0.15$, the firm's $\left(\beta_{iL}^v - \overline{s}_{iL}^v\right)$ gap is 0.44. This firm reallocates 0.44×0.15 = 0.07 of its labour to more production efficient firms and that has productivity-enhancing effects on APG. Second, a negative value can arise for a low-cost $(MC_{1L} < MC_{2L})$ downsizing $(\ln X_{imt} < \ln X_{imt-1})$ firm 1 relative to firm 2, due to firm 1's investment in technological innovation either in the presence of weak demand for its products or substantial price inelasticity of demand faced by the firm. In this case, higher productivity firm 1 reallocates labour to lower productivity firm 2.

The Estimating Equation has two key elements that require particular attention: the regression residual and consistent estimates of revenue-function elasticities. Since this is a value-added function, only labour and capital inputs are relevant: Working Proprietors $(L_i^{WP})$, Paid Employees $(L_i^{PE})$, and Plant, Machinery and Equipment (PME) $(K_i^{PME})$. This reduces to Eq. 6 as follows

$$ln\,\Omega_i^v = ln\,V\,A_i - (\beta_{iPE}^v L_i^{PE} + \beta_{iWP}^v L_i^{WP} + \beta_{iPME}^v K_i^{PME}). \quad [6]$$

The consistent parametric estimation of production technologies in Eq. 6 is performed separately using the control function approach. The idea is to estimate a Cobb-Douglas revenue function to recover value-added output elasticities and the residual for use in the computation of within-plant productivity, factor-input reallocation, and net-entry in the Estimating Equation. Variable definitions and data issues for the Estimating Equation are covered in Section 4.

## 4. The Data and Preliminary Results

This section presents and discusses variable definitions, data sources, and representativeness of the survey data. The dissection of the structural APG decomposition into entry/exit margins, productivity, and input reallocation particularly invites thinking about their impact on relevant markets as used in antitrust parlance (Baker, 2007). Notwithstanding the inadequacy of Fourth-Digit ISIC classifications as measures of economic markets; for instance, we deem the component parts of the decomposition to apply to these industries as distinct, relevant markets.

### 4.1. Definition of Variables

**Employment ($L_{it}$)**: There are two main categories of employment. These are the number of Paid Employees (PE) and Working Proprietors (WP).

**Intermediate Inputs ($M_{ijt}$)**: Intermediate inputs ideally refer to expenditure in Material inputs as well as Electricity, Water, and Fuel. The last three variables are aggregated and reported as Energy. In this paper, $M_{ijt}$ refers to Material inputs produced by firm $i$ for consumption either by firm $i$ or firm $j$ or both firms at time $t$.

**Investment ($I_{it}$)**: Firm $i$'s investment expenditure at time $t$ is measured as $I_{it} = Exp_{it} - Ret_{it}$, where $Exp_{it}$ is firm $i$'s value of expenditure on fixed capital investment and $Ret_{it}$ is the retirement value of fixed capital investment.

**Capital Stock ($K_{it}$)**: The capital series measurement is based on the Perpetual Inventory Method (PIM) which takes a first-order Markov chain specification; i.e., $K_{it} = I_{it} + (1-\delta)K_{it-1}$. To obviate the difficulties posed by initial conditions problems, the initial values of the capital variable are measured as

$K_0 = \frac{GFK_0}{\delta + g_{GFK}}$, where $K_0$ is the initial PME capital stock, $GFK_0$ is 1994 gross fixed capital formation of PME, and $g_{GFK}$ is the growth rate of PME.[12] We 'winsorized' capital stock to trim one percent of firms at the tails of the capital kernel distribution in order to minimize administrative recording errors. There is generally no lumpy investment during the sample period, except the large transaction between an upstream plant and downstream firm in 1998/1999.

**Entry ($E_{it}$)**: A new entrant is a firm that is not present in the database at time $t-1$ but present at time $t$.

**Exit ($\chi_{it}$)**: An exiting firm is one that is present at time $t-1$ in the database and absent at time $t$.

**Incumbent ($C_{it}$)**: An incumbent plant is one that is present at time $t-1$ and also present at time $t$.

**Salaries and Wages ($W_{it}$)**: Personnel expenditure refers to separate costs of Paid Employees and Working Proprietors.

**Revenue ($Rev_{it}$)**: Revenue refers to export ($ichrprx_{it}$) and domestic ($ichrprl_{it}$) sales revenue.

**Deflators ($Def_t$)**: There are three deflators used in this study. We use the $MVADEF_t$ sourced from the World Bank Indicators (WBI) developed for the deflation of Manufacturing $Rev_{it}$ and Material inputs. The capital stock series is also deflated using the WBI gross fixed capital formation deflator while we deflate Salaries and Wages using the Consumer Price Index($CPI_t$).

### 4.2. Data Sources

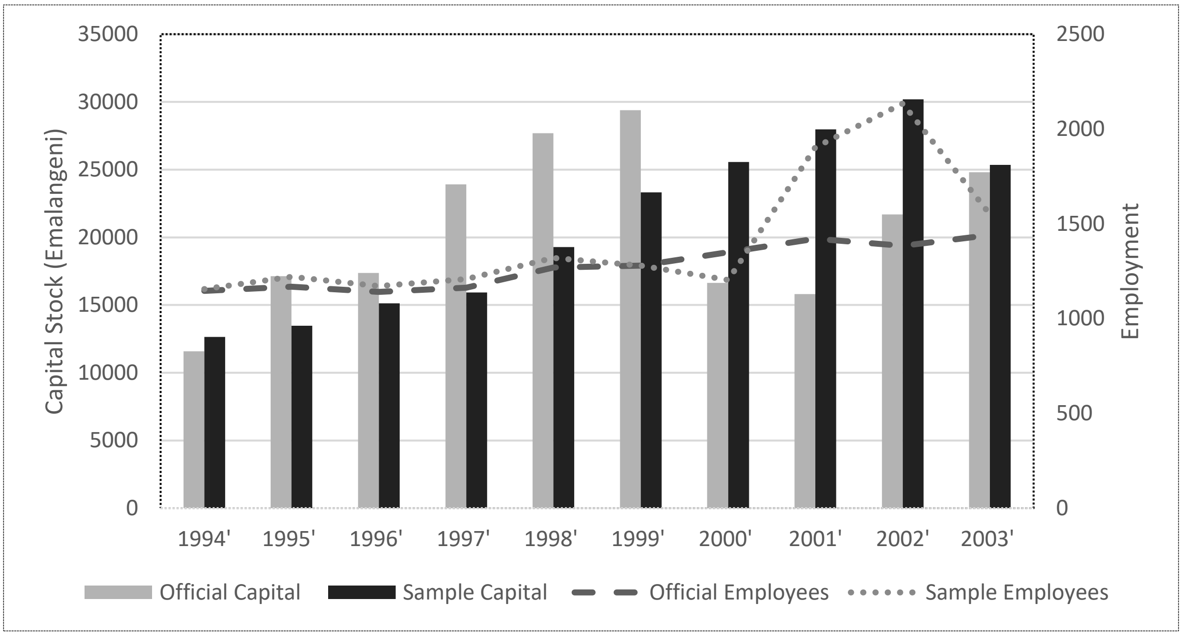
The dataset came in three different files from the Central Statistical Office (CSO) of Eswatini. There was one database for large firms, one for small firms, and one for fixed capital investments. The three databases were connected through firm identity codes and contained financial information and related variables. In other words, one file contained information on firms whose contribution in their industry exceeded a certain threshold determined for that year. A second file contained information on firms whose contribution fell below the said threshold. There was no variation in firm identities regardless of periodic shifts across threshold. Whether a firm is above or below the threshold is inconsequential because both firm types remain in the overall database and can be tracked using their identity codes. The final dataset contained information about investments, including retirement revenues and expenditure on capital investments. Data cleaning and selection entailed exclusions involving zero, negative or missing employment, revenue, and/or material inputs.[13] The composite dataset used for analysis had 2 179 plant-year records and 335 firms that ever engaged in production in the 1994-2003 period.

---

[12] See Hall and Jones (1999), footnote 5, for handling the initial conditions problem.
[13] See Gopinath et al. (2017) for the case of southern Europe.

### 4.3. Data Representativeness

Representativeness of the census data is gauged through comparison with official data. Figure 4 presents a graph of inputs contained in the census dataset and official data published in International Monetary Fund (IMF) annual reports. In the period 1994-1999, the official record of fixed capital stock shows relatively rapid growth compared to census capital represented by the stock of PME. This variable declined sharply in 2000 only to rise again in 2002 in response to admission of Swazi Textile and Apparel industries to the African Growth and Opportunity Act (AGOA). These merchandize exports to the US accounted for 4% of total exports since inception of the Act. In contrast, the census data shows a steady growth only to converge to official fixed capital aggregates in 2003. A similar trend is observed in the case of paid employees; *albeit*, with census employment overshooting its counterpart official employment series. This behaviour of employment is also consistent with the entry of new labour-intensive firms in the Textile and Apparel from 2001.




***Note:*** **Capital stock series is expressed in constant 2000 prices.**

**Figure 4: Distribution of Official and Sample Factor-Inputs.**

**Source: Author using Official Macroeconomic Indicators come from the IMF Annual Country Reports (1999, 2000, 2003 and 2006) for real capital stock and employment.**

The 1994-2003 sample period coincides with a trade liberalization episode in SACU precipitated by the re-introduction of the South Africa back to the world economy. The new and tougher competition presented by the entry of more productive plants in the Customs Union forced the exit of some domestic plants. Local firms further experienced supply-side constraints arising from; *inter alia*, distortionary effects of government involvement, un-competitive investment environment, and regulatory restrictions. As a results, economic growth measured by GDP fell from eight percentage points in the 1980s to 3.8 percentage points in the 1990s and 2.3 percentage points in the 2000s. This pattern of growth mirrors FDI inflows from 5.83

percentage points in the 1980s to 2.5 percentage points in the 1990s. Moreover, Eswatini has been losing skilled labour to South Africa since inception of trade reforms (Edwards, et al., 2013).

## 5. Main Results

This section reports the main results on the Estimating Equation under price-taking conditions and product homogeneity assumptions so that there is no distinction between revenue function and physical quantity elasticities in subsection 5.1. These assumptions provide baseline results for comparison of reallocation and within-firm productivity with findings in the empirical APG literature. On the other hand, assumptions about monopolistic competition in subsection 5.2 clarify the impact of adjustments in the Estimating Equation and effects of controlling for supply and demand factors.

### 5.1. Baseline Results under Price-Taking and Product Homogeneity Assumptions

The key assumption in this subsection is that plants take prices as given and products are homogeneous within industries. In the absence of factor-augmenting technical change, plants rely on Cobb-Douglas production technology to help simplify estimation of revenue elasticities with respect to factor inputs. We apply proxy methods to each of the 13 Two-Digit ISIC industry comprising 49 Four-Digit ISIC industries in the sample period.

Table 1 summarises the results under price-taking to enable direct comparison with the current practice in the literature. It therefore decomposes APG into annual total factor productivity, primary input reallocation and net entry effects. The second and third columns show changes in real value added and APG, respectively. It is striking to observe the degree of precision in APG tracking value-added output growth. On average, the sector reports an estimated annual real value-added growth rate of 54.59 percent, and 54.54 percent of that growth comes from APG.

More importantly, the baseline results relating to productivity and reallocation surprisingly mimic findings by Nishida et al. (2014) for Chile, Colombia and Slovenia. For instance, note that the average (median) total factor productivity is-3.61 (-2.69) percent per year compared to an average of 0.95 percent for Chile, 0.25 percent for Columbia and 2.17 percent for Slovenia generated using the similar methods. In South Africa, Aghion et al. (2008) reports productivity growth of 0.04. Within-plant productivity growth in Eswatini is positive in only four out of the nine years. As a result, the observed patterns of annual productivity and related moments suggest that the sector experienced persistent declines in year-on-year plant-level total factor productivity. This poor performance as reflected in the first-moments of multifactor productivity is concerning given that this measure has long-run economic growth effects.

Furthermore, looking at input reallocation provides insights into cross-firm movements of economic activity among incumbent plants. The structural APG decomposition identifies resource movements from low to high (VMP-MC) gap plants, and vice versa. The outcome of input churning is a positive average value of 0.15 percent in multifactor reallocation thereby increasing APG. Clearly, the overall average reallocation compares with 1.60 percent for Chile, 3.63 percent for Columbia and 3.42 percent for Slovenia as reported in Nishida *et al.* (2014). [14]

**Table 1: Baseline APG Decomposition under Price-Taking Conditions for 49 Four-Digit ISIC Industries**

| | | | APG Decomposition: (0) = (1) + (2) + (3) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Reallocation | | | | |
| Year | Value-Added Growth | APG (0) | Total Factor Productivity (1) | Total Factor Reallocation (2) | Capital Reallocation | Total Labour Reallocation | Paid Employees' Reallocation | Net Entry (3) |
| 1995 | 7.76 | 7.71 | -4.43 | -4.31 | -13.07 | 8.76 | 3.15 | 16.45 |
| 1996 | 23.10 | 23.03 | 2.27 | -6.98 | -8.19 | 1.21 | 1.51 | 27.75 |
| 1997 | -44.35 | -44.25 | -2.69 | 18.13 | 8.29 | 9.84 | 9.92 | -59.69 |
| 1998 | 265.55 | 265.30 | 2.31 | -2.38 | -2.48 | 0.10 | 0.07 | 265.37 |
| 1999 | 275.57 | 275.42 | 0.64 | 9.81 | 6.8 | 3.01 | 3.03 | 264.97 |
| 2000 | -16.28 | -16.27 | -15.30 | -5.16 | -5.77 | 0.61 | 0.74 | 4.18 |
| 2001 | 37.42 | 37.39 | 9.03 | 20.10 | 20.18 | -0.08 | 0.03 | 8.25 |
| 2002 | -20.74 | -20.75 | -3.56 | -29.01 | -27.65 | -1.36 | 0.30 | 11.82 |
| 2003 | -36.71 | -36.67 | -20.74 | 1.12 | -6.02 | 7.14 | 9.75 | -17.05 |
| **Mean** | **54.59** | **54.54** | **-3.61** | **0.15** | **-3.10** | **3.25** | **3.17** | **58.01** |
| **Median** | **7.76** | **7.71** | **-2.69** | **-2.38** | **-3.59** | **1.21** | **1.51** | **11.82** |
| **Std Dev** | **125.32** | **125.23** | **9.21** | **14.88** | **10.66** | **4.22** | **3.96** | **120.14** |

*Notes*: Firms are price-takers in markets for product varieties assumed homogeneous within industries. Hence, revenue function elasticities equate to physical output elasticities. Numbers in cells are percentage points. The plant-level multifactor productivity uses production function parameters that vary across the 2-digit ISIC code obtained by using Wooldridge (2009); i.e., ivreg29. APG represents aggregate productivity growth rate with entry and exit, and is the aggregate change in final demand *minus* aggregate change in expenditure in inputs, holding inputs constant. Value-added output shares (Domar) are weights. APG decomposes into four components, excluding reallocation of Working Proprietors: (1) Total Factor Productivity, (2) reallocation, and (3) net-entry term, using the Estimating Equation, Eq. 5 in text.

**Source**: Author's Calculations.

The general pattern of factor-input reallocation is strongly correlated with reshuffling patterns exhibited by the direction of change in 'Paid Employee Reallocation', and shows cross-plant reshuffling in the productivity-enhancing direction. Average reallocation is 3.17 percent per year and consists labour reallocation from low to high productivity plants. To add more clarity on this, we isolate labour reallocation from the contribution of all inputs put together. This produces 3.25 percent as the average annual rate of labour reallocation, and we report only two instances of negative reallocation out of the nine years studied. Paid employment shows growth in every year and accounts for an average of about 98 percent of all labour reallocated per year. However, the magnitude of paid labour reallocation diminishes with time; *albeit*, with infrequent and random positive shocks. On the other hand, fixed capital stock reallocates from high-productivity to under-performing plants, regardless of controls on productivity dispersion.

[14] An interesting fact about the Nishida et al. (2014) results is that the factor-input reallocation component of APG dominates within-firm effects when the structural accounting decomposition is used, and the converse is true with using statistical accounting decomposition methods of industry productivity growth.

Thus, the analysis reveals that the contribution by the labour reallocation growth to APG *dominates* 'Within-Firm' effects. Firms were not investing more in improving production efficiency through innovation and adoption of new technologies as well as human capabilities than they were moving labour to higher activity producers. On the other hand, the extensive margin of cross-firm reallocation of inputs accounts for most of the changes observed in APG. The annual average of net entry contribution to APG was 58 percent due to the dramatic increase of APG in 1998 and 1999. This pattern of high contribution by net entry is consistent with extensive margin effects of trade liberalization that increase opportunities for mergers and acquisitions as well as business restructuring and retrenchments.

The productivity losses experienced within-firms and the near-zero factor-input reallocation effect on APG are both concerning for policymakers. The real issue is; however, whether or not price-taking is a realistic supposition for production and consumption in manufacturing. Evidence abounds that firms do charge prices higher than the hard-to-observe marginal costs (see, for example, De Loecker et al., 2016; De Loecker et al., 2020; Mhlanga and Rankin, 2021). It seems sensible therefore to consider results from price-taking assumptions as a reflection of lower bounds for measures of productivity and input reallocation. Moreover, as is well known from Foster et al. (2016), $lnTFP^{rr}$is a function of $lnTFPQ$ and idiosyncratic demand shifters under isoelastic demand conditions. The related analysis is carried out in the next section to recover the effects of monopolistic competition on APG.

### 5.2. Monopolistic Competition with Heterogeneous Markups

This section considers price-setting market conditions in capturing the effects of variable pricing power across plants, given that markup pricing is an integral part of static allocative efficiency (Peters, 2020). It also considers several relationships involving the demand elasticity and the shape of the demand function, also known as the demand manifold (see Mrázová and Neary, 2017). The demand manifold determines independent relationships between itself and characteristics of firms and the HARA utility function. Second, the demand parameter scales $lnTFP^{rr}$ and $\beta_{im}^{v}$ to back out $\alpha_{im}^{v}$ in order to recover $lnTFPQ$.

**$lnTFP^{rr}$ and Fundamentals**: The standard inverse residual demand specification is presented as $P_i = P_s \left(\frac{Q_s}{Q_i}\right)^{1-\rho} \xi_i$, where $P_i$ and $P_s$ are respective firm-level and aggregate industry prices, $Q_i$ is quantity produced by firm $i$, $Q_s$ is quantity produced in sector $s$, $\xi_i$ is an idiosyncratic demand shifter, and $\rho \in (0,1)$ denotes a demand parameter that determines the elasticity of demand (cf Foster et al., 2017). Thus, the constant returns to scale (CRS) production technology adopted here implies the revenue functional form $P_iQ_i = P_s \left(\frac{Q_s}{Q_i}\right)^{1-\rho} Q_i\xi_i$. This function can be expressed as $P_iQ_i = P_sQ_s^{1-\rho}Q_i^{\rho-1}Q_i\xi_i = P_sQ_s^{1-\rho}Q_i^{\rho}\xi_i$. Taking logarithms on both sides, the revenue function that produces the conceptual log Total Factor Productivity Residual ($lnTFPR$) for firm $i$ becomes:

$$p_i + q_i = p_s + (1-\rho)q_s + \rho q_i + ln\,\xi_i$$

$$= p_s + (1-\rho)q_s + \rho(\sum_{m=1}^{M} \alpha_{im}^{v} x_{im} + a_i) + ln\, \xi_i \quad [7]$$

where $\alpha_{im}^{v}$ represents factor $m$ elasticity of value-added based physical output produced by firm $i$, $a_i = lnTFPQ_i$ is technical efficiency and $x_{im}$ denotes the log of factor-input for firm $i$. At this point, it is clear from Eq.8 that $\rho\alpha_{im}^{v} = \beta_{im}^{v}$. We can thus show that the characterization of $lnTFP$ based on estimation under isoelastic demand is $lnTFP^{rr} = \rho \times lnTFPQ$. In this setting, $lnTFP^{rr}$ rises if a firm experiences a positive demand shift and/or a technical efficiency boost, holding output elasticity of demand and aggregate prices constant.

**HARA Preferences**: The additively separable HARA utility function faced by heterogeneous firms is one of the commonly used preferences according to Dhingra and Morrow (2019). In its general form, the preferred specification of the HARA utility is

$$u(Q_{Hi}) = \xi_{Hi} \times \left[\left(\frac{Q_{Hi}}{1-\rho_{Hi}} + \gamma_{Hi}\right)^{\rho_{Hi}} - \gamma_{Hi}{}^{\rho}\right] \times \left(\frac{\rho_{Hi}}{1-\rho_{Hi}}\right)^{-1},$$

that is strictly increasing and concave in quantity consumed, where $\xi_{Hi}$ is a plant-specific demand shifter, the unit-free demand manifold statistics specified as $\{\varepsilon_{DHi}, \gamma_{Hi}\}$. The demand parameter restrictions that can also be estimated are $\gamma_{Hi} > \frac{q_{Hi}}{\rho_{Hi}-1}$ and $\rho_{Hi} \in (0,1)$, and the related HARA utility function yields APG results that are consistent with fundamental economic reasoning.[15] There are three scenarios of interest concerning the parameter values of the utility function. First, if $\rho_{Hi} = 0$, then $\gamma_{Hi} > -q_{Hi}$. The relationship with negative quantity demanded in this case produces misaligned preferences that have neither *a priori* nor empirical appeal (Dhingra and Morrow, 2019). Second, with $\gamma_{Hi} = 0$, preferences reduce to the CES utility function which purges all pro-competitive effects in the market due to markup and demand invariance to shocks because of the common markup $\mu_{CES} = \rho_{CES}^{-1}$. Third, the preference technology reflecting $\gamma_{Hi} > 0$ is the main focus of this paper, where the inverse residual demand curve for differentiated products is given by

$$P_{Hi} = u'(Q_{Hi}) = \xi_i \times \left(\frac{Q_{Hi}}{1-\rho_{Hi}} + \gamma_{Hi}\right)^{\rho_{Hi}-1}.$$

From this preference structure, with $\varepsilon_{DHi} = \rho_{Hi} - 1$ as the inverse price elasticity of residual demand, profit maximization under monopolistic competition leads to variable markups of the form

$$\frac{P_{Hi}}{MC_{Hi}} = 1 + \mu_{Hi} = \frac{u\prime(Q_{Hi})}{u\prime(Q_{Hi}) + u\prime\prime(Q_{Hi})Q_{Hi}} = \frac{Q_{Hi} + \gamma_{Hi}(1-\rho_{Hi})}{\rho_{Hi}Q_i + \gamma_{Hi}(1-\rho_{Hi})}$$

[15] See Perets and Yashiv (2015) who assert that the HARA demand structure is not only useful because of its tractability, but it arises from fundamental economic motivation. Furthermore, Pollak (1971) characterizes HARA preferences as the only utility function that is consistent with both additive separability and quasi-homotheticity.

henceforth referred to as the HARA Markup Equation which obeys Marshal's Second Law of Demand (MSLD), and $MC_{Hi}$ is the marginal cost for firm $i$. MSLD posits that the price elasticity of residual demand increases with price and, by extension, markups (Krugman, 1979; Zhelobokdo et al., 2012; and Mrázová and Neary, 2017). Moreover, the *locus classicus* for the behaviour of the demand manifold as a sufficient statistic for many comparative statics' predictions is Mrázová and Neary (2017).

The measurement of $lnTFPQ$ and $\alpha_{im}^{v}$ requires estimation of $\rho_{Hi}$ to scale $lnTFP^{rr}$ and revenue elasticities. There are at least three ways for estimating $\rho_{Hi}$. First, there exists a method based on Klette and Grilitches (1996) that undertakes joint computation of $\beta_{im}^{v}$ and $\rho_{Hi}$ by including industry output in a regression of Eq. 8 to back out $\alpha_{im}^{v}$. Second, another method estimates $\rho_{Hi}$ directly from Table 2 as a function of idiosyncratic markups, a measure of convexity of the demand system, and industry output. Finally, another method estimates $\alpha_{im}^{v}$ using factor-input expenditure shares.[16]

Table 2 presents a closed-form structure of HARA parameters, and its special Constant Elasticity of Substitution (CES) case, to scale the APG decomposition under price-setting conditions. Notably, the characterization of product markets with a CES demand function implies that the elasticity of substitution is constant and common across all product varieties as in column 3. A few more drawbacks of the CES market structure are contained and explained in Zhelobokdo, et al. (2012). The analysis here rather relies on common average markups to derive the demand parameter to scale $lnTFP^{rr}$ and revenue elasticities. In contrast, the characterization of the HARA economy presents the variable HARA Markup Equation as a function of firm- and product-specific output and determinants of the demand manifold. Similarly, the last column provides an expression for the HARA scaling factor, $\rho_{Hi}$, as a function of firm- and product-specific output, markups, and a demand parameter. It also presents an expression for the curvature of the demand system, $\gamma_{Hi}$.

**Table 2: The Closed-Form Structure of Markups and Demand Manifold**

| **Market Conditions** | **Markup Equation, $1+\mu$** | **Demand Parameters, $\rho$ and $\gamma$** |
|---|---|---|
| CES | $\mu_{CES} = \rho_{CES}^{-1}$ | $\rho_{CES} = \mu_{CES}^{-1}$ |
| HARA | $1+\mu_{Hi} = \dfrac{Q_{Hi} + \gamma_{Hi}(1-\rho_{Hi})}{\rho_{Hi}Q_{Hi} + \gamma_{Hi}(1-\rho_{Hi})}$ | $\rho_{Hi} = \dfrac{Q_{Hi} - \gamma_{Hi}\mu_{Hi}}{Q_{Hi} - \gamma_{Hi}\mu_{i} + Q_{Hi}\mu_{Hi}}$ |
| HARA | | $\gamma_{Hi} = \dfrac{Q_{Hi} - \rho_{Hi}Q_{Hi}(1-\mu_{Hi})}{\mu_{Hi}(1-\rho_{Hi})}$ |

**Source**: Author.

[16] The analysis here experimented with all three approaches and found that the Klette-Grilitches method indeed pushes the data too hard while input expenditure shares produced coefficients that are significantly different from both methods and may be representing long-run estimates of $\alpha_{im}^{v}$, as in Foster et al. (2017).

To fix ideas, a back-of-the-envelope calculation setting $\gamma_{Hi} = 26$ that obtains in the US and $\rho_{Hi} = 0.5$, defining the inverse price elasticity of demand as $\rho_{Hi} - 1 = -0.5$ and choosing physical quantity as $Q_{Hi} = 500kg$ yield a markup of $(1 + \mu_{Hi})$ =1.95. This proof-of-concept example incidentally compares well with Foster et al. (2017) who found in real-world panel data an average markup of 2.08, with an implied average demand parameter of $\rho = 0.48$ that is common to all industries.[17] The intuition here is that the HARA utility delivers plausible results if measurement of the curvature of the demand function and the demand elasticity is precise.

### 5.2.1. Empirics of Monotone Comparative Statics in the HARA Economy

In this analysis, producer and/or industry level variability in markups and output pins down variations in the demand parameters to similar levels of granularity as shown in Table 2. As a result, the data-driven distribution of the demand manifold is $\{(\varepsilon_{DHi}, \gamma_{Hi}) : \varepsilon_{DHi} \in (-31.2, 10.5)\ and\ \gamma_{Hi} \in (2.08, 9.75)\ \}$.[18] There is an important observation regarding the characterization of the data-driven demand manifold: the demand elasticity undershoots the theoretical lower bound of $\varepsilon_{DHi}$ =1 to $\varepsilon_{DHi}$ =-31.2 while the shape parameter space lies beyond its own theoretical space bounded above as $\gamma_{Hi}$ =2.

Turning to relationships between the demand manifold and producer idiosyncrasies, the statistics relate to plant performance and economic outcomes. The literature holds the notion that high markup plants are technically efficient and are in high capital intensity industries. According to this view, producers that are heterogeneous in productivity self-select to charge higher markups in higher capital/labour ratio industries. This is evident in Panels A and B in Figure 5 that relate technical efficiency to markups and capital intensity. For instance, the cross-sectional dispersion in technical efficiency grows with markups, indicating the coexistence of inefficient and efficient firms (cf. Mrázová et al., 2021). As Foster et al. (2018) points out, higher technical efficiency plants in higher markup industries are generally more capital intensive, and that technical efficiency reflects product quality than process efficiency.

In Panel C, the elasticity of demand relates to markups in the reference period. The elasticity of demand is largely an increasing function of markups. We further observe an inverted U-shape relationship between the demand elasticity and industry output in Panel D. This means the elasticity of demand increases with industry output up to an inflection point and then declines with a further increase in industry output. In a

---

[17] Haltiwanger et al. (2018) estimate the product-specific $\gamma_{Hi}$ by regressing log prices on log quantities and inverse quantities in levels, and rely on the proportionality relationship between $\gamma_{Hi}$ and $Q_{Hi_i}$; i.e., a choice of any factor of $Q_{Hi}$ changes $\gamma_{Hi}$ by the same factor.

[18] However, the admissible region for the demand manifold is *a priori* $\{(\varepsilon, \gamma) : \varepsilon \in (1, \infty)\ and\ \gamma \in (-\infty, 2)\ \}$, but the evidence cited in Mrázová and Neary (2017) is that $\{(\varepsilon, \gamma) : \varepsilon \in (1, 4.5)\ and\ \gamma \in (-2, 2)\ \}$. In contrast, $\gamma = 26$ for the US in Foster et al. (2018), which is much higher than the empirical upper bound of $\gamma = 2$. In De Loecker et al. (2016), $\mu$ =0.34 and the pass-through coefficient $\varkappa$=0.305 imply $\varepsilon$ =3.941 and $\gamma \in (-3.317, 0.411)$. But; of course, our goal is to measure and/or estimate rather than test any theory.

similar vein, Burya and Mishra (2022) found higher elasticity of demand for smaller producers and that the demand elasticity depended on market share. Therefore, the demand elasticity increases with sales *a priori* (alternatively, declines with price) if, and only if, the inverse demand function is superconvex (more convex than CES).[19] Interpreted in the context of Mrázová and Neary (2017), $\varepsilon_{DHi}$ in Panel D decreases with industry output subject to a subconvex (less convex than CES) inverse demand function to the right of the inflection point and increases with industry output in relation to a superconvex inverse demand function to the left of the inflection point. On the other hand, larger firms also experienced progressively depressed markups on account of the South African trade liberalization episode which spilled over to the rest of the Customs Union.

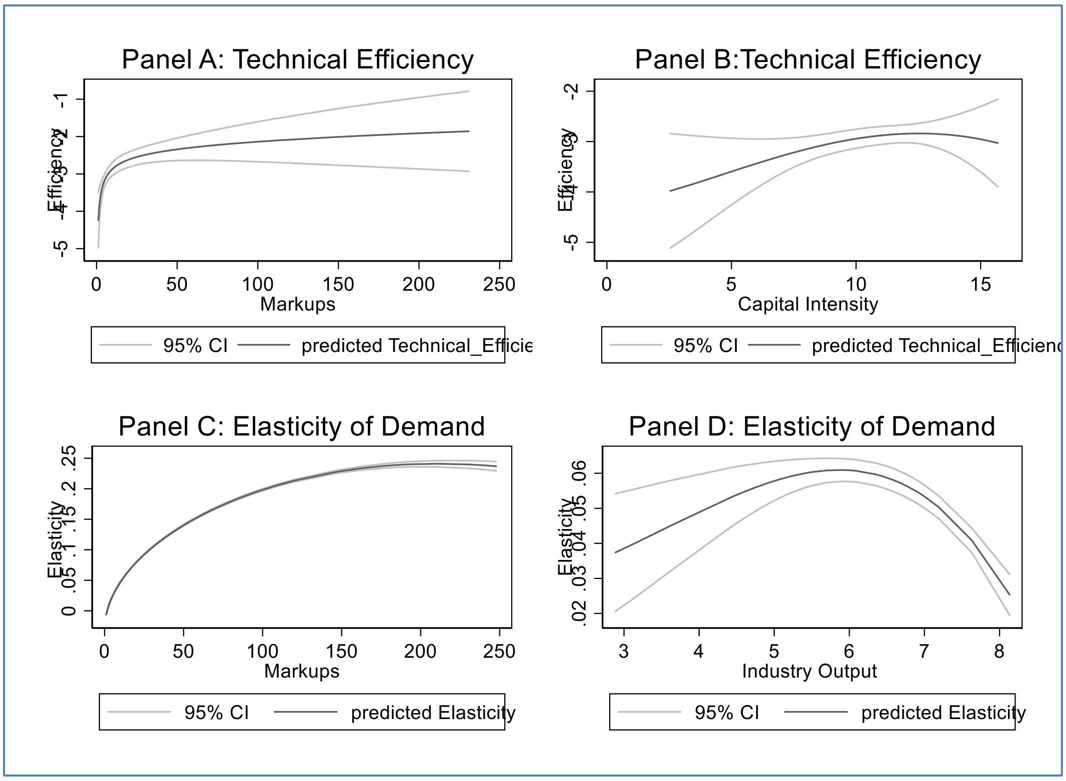



**Figure 5: Comparative Statics Predictions.**

**Source:** Author.

### 5.2.2. Benchmark Results under Monopolistic Competition

This subsection provides key benchmark results under monopolistic competition assumptions against which controlled environments are measured and evaluated. Given the utility technology and an associated inverse demand functional form, we adopt deflated industry-output as a proxy for the quantity variable. Then two measures of markup pricing are considered: 1) common markups relevant for CES preferences

[19] Mrázová and Neary (2017) define superconvexity as an arbitrary point in the $\{\varepsilon_{DHi}, \gamma_{Hi}\}$ space if it is more convex at that point than the CES demand function with the same elasticity. In contrast, if it is less convex than the CES demand function, then it is subconvex.

are computed in Lerner index form as $\mu_i = \frac{P_iQ_i - C_iQ_i}{C_iQ_i}$ instead of using the production approach due to data sparsity constraints which render revenue elasticities negative, zero, or even inestimable for some firms. 2) variable markups within industries are defined by the Markup Equation, along with expressions for demand function convexity and the scaling demand parameter, in Table 2 under HARA preferences. The focus of analysis here is the impact of the scaling demand factor, $\rho$, on within-firm productivity and input reallocation as well as on the related measured dispersion.

**Table 3: Benchmark Productivity and Reallocation under CES and HARA Markups for 49 Four-Digit ISIC Industries**

| PANEL A: Constant Markups; $\boldsymbol{\rho_{CES} = (1+\mu_{CES})^{-1}}$, where $\boldsymbol{\mu_{CES} = 0.62}$ | | | | | |
|---|---|---|---|---|---|
| **Year** | **Technical Efficiency** | **Multifactor Reallocation** | **Total Labour Reallocation** | **Paid Employee Reallocation** | **Capital Reallocation** |
| 1995 | -5.73 | 0.25 | 2.10 | 2.74 | -1.84 |
| 1996 | 1.31 | 3.79 | 0.45 | 0.62 | 3.34 |
| 1997 | -2.41 | 11.92 | 13.52 | 13.52 | -1.60 |
| 1998 | 0.39 | 8.99 | 1.89 | 1.89 | 7.09 |
| 1999 | 1.65 | 21.37 | 7.16 | 7.16 | 14.20 |
| 2000 | -47.66 | -7.10 | 0.79 | 0.74 | -7.90 |
| 2001 | 22.48 | 13.25 | -0.34 | -0.76 | 13.60 |
| 2002 | -19.05 | -1.48 | 0.43 | 2.85 | -1.92 |
| 2003 | -11.02 | -0.52 | 2.97 | 1.93 | -3.49 |
| **Mean** | **-6.00** | **5.04** | **2.90** | **3.07** | **2.14** |
| **Median** | **-1.20** | **2.02** | **1.34** | **1.91** | **-0.80** |
| **Std Dev** | **18.08** | **8.61** | **4.32** | **4.27** | **7.35** |
| **PANEL B: Heterogeneous Markups;** $\rho_{Hi} = \frac{Q_{Hi} - \gamma_{Hi}\mu_{Hi}}{Q_{Hi} - \gamma_{Hi}\mu_i + Q_{Hi}\mu_{Hi}}$ **and** $\gamma_{Hi} = $ **5.92** | | | | | |
| **Year** | **Technical Efficiency** | **Multifactor Reallocation** | **Total Labour Reallocation** | **Paid Employee Reallocation** | **Capital Reallocation** |
| 1995 | -11.46 | 2.73 | 3.81 | 6.76 | -1.09 |
| 1996 | 2.68 | -6.98 | -1.86 | -0.81 | -5.12 |
| 1997 | -4.80 | 19.62 | 23.41 | 23.41 | -3.78 |
| 1998 | 0.74 | -0.85 | 2.65 | 2.65 | -3.50 |
| 1999 | 2.66 | 12.94 | 19.19 | 19.19 | -6.25 |
| 2000 | -89.55 | 52.76 | 0.55 | 0.54 | 52.21 |
| 2001 | 41.45 | 2.01 | -0.80 | -1.71 | 2.81 |
| 2002 | -35.17 | -9.98 | -4.75 | 8.20 | -5.22 |
| 2003 | -20.69 | 7.93 | 4.71 | 1.51 | 3.22 |
| **Mean** | **-11.41** | **8.91** | **4.69** | **5.97** | **3.69** |
| **Median** | **-2.40** | **2.72** | **1.60** | **2.08** | **-3.50** |
| **Std Dev** | **33.78** | **18.86** | **9.23** | **8.73** | **18.51** |

*Notes*: Producers are price-setters with product differentiation and markup heterogeneity. Scaling revenue function residual and elasticities by the demand parameter equates them to technical efficiency and physical output elasticities, respectively. Numbers in cells are percentage points. The plant-level multifactor productivity uses production function parameters that vary across the 2-digit ISIC code obtained by using Wooldridge (2009); i.e., ivreg29. APG represents aggregate productivity growth rate with entry and exit, and is the aggregate change in final demand *minus* aggregate change in expenditure in inputs, holding inputs constant. Value-added output shares (Domar) are weights. APG decomposes into: technical efficiency and reallocation using the Estimating Equation, Eq. 5 in the text.

**Source**: Author's Calculations.

Under CES markets in Panel A of Table 3, the common markup is based on Mhlanga and Rankin (2021) who use the same dataset to measure $\mu_{CES} = 0.62$[20]. First and foremost, scaling $lnTFP^{rr}$ by $\rho_{CES}$ amplifies within-plant productivity a 1.7-fold in absolute value. It also amplifies the standard deviation measure of dispersion a 1.1-fold. In the case of factor-input reallocation, the scaling factor transforms reallocation losses in the capital stock variable under price-taking environments to reallocation gains in CES markets. It also exhibits moderate attenuation of Paid Employment reallocation. All components of resource reallocation, but capital stock, exhibit amplified dispersion.

Furthermore, idiosyncratic markup heterogeneity under HARA preferences in Panel B, with the shape parameter of the demand function at its midpoint level of 5.92, provides similar patterns of change as in the CES case. However, these have somewhat pronounced orders of magnitude under HARA utility assumptions. Such formulation transforms the first moments of price-taking outcomes from $(lnTFP^{rr}, lnTFR,\ lnE) \in (-3.61, 0.15, 3.17)$ to HARA outcomes $(lnTFPQ, lnTFR,\ lnE) \in (-11.41, 8.91, 5.97)$, where $lnTFR$ is log multifactor reallocation, and $lnE$ is log Paid Employment. That is, $\rho_{HARA}$ increased $lnTFP^{rr}$ a 3.2-fold under HARA market assumptions while significantly increasing multifactor reallocation a circa 60-fold, with Paid Employment making the highest contribution to this change. The measured dispersion also exhibited significant magnification effects of the scaling factor.

The observed productivity and reallocation patterns under HARA economies are an outcome of at least three characteristics of consumer demand: 1) the extent of noise in $\rho_{Hi}$, 2) the extent of idiosyncratic variability in $\rho_{Hi}$, and 3) the interaction between $\rho_{Hi}$ and $lnTFP^{rr}$, and between $\rho_{Hi}$ and factor-input reallocation. Nonetheless, although our results are implemented with firm-specific markups within industries, the impact of controlling for demand characteristics in the estimation of $lnTFPQ$ and factor-input reallocation mimics the more aggregated results in Foster et al. (2017).

Thus, monopolistic competition with either a CES or HARA demand system had *a fortiori* marked magnification effects on production inefficiency and multifactor reallocation in the manufacturing sector of Eswatini during the trade liberalization period in SACU.[21]

#### 5.2.3. Robustness Checks for Productivity and Reallocation under HARA Markups

[20] Foster et al. (2017) fixes a common markup for U.S. data at $\mu$ =10% and $\mu$ =25% corresponding to $\rho$ =0.9 and $\rho$ =0.8.

[21] However, caution must be taken when interpreting the results under the HARA economy. The non-observability of prices and physical-output renders the firm-level variability of markups and therefore the parameter that determines the elasticity of demand very noisy. Not surprisingly, scaling $TFP^{rr}$ to recover TFPQ and scaling the revenue-function elasticities to recover physical-output elasticities make these terms explode (Foster et al., 2017). Thus, the analysis calls for precision in the estimation of demand parameters, given that a revenue-function does indirectly produce estimates of physical-output elasticities. However, there is no cause for alarm in the present analysis, particularly if extreme values are controlled for.

This subsection explores whether controlling for structural features of the demand function and plant characteristics generate extraneous variation in technical efficiency and factor-input reallocation. In other words, does varying $\gamma_{Hi}$ or demand function convexity or firms' downsizing/upsizing produce first-order and second-order moments that are insignificantly at variance with benchmark outcomes, *ceteris paribus*? Subsection 5.2.3 A. examine the sensitivity of the benchmark results to the curvature of the demand function using its lower and upper bounds. In subsection 5.2.3 B., we examine the robustness of our benchmark results to market distinction in terms of superconvexity and subconvexity of demand while subsection 5.2.3 C. looks at the results' sensitivity to plant expansion and contraction. Subsection 5.2.3 D. documents demand elasticities and markups by productivity, employment, and production.

**A. Variation in the Curvature of the Demand Function**

The assumption of monopolistic competition suggests that firms are without supply curves but rather their adjustment occurs along the marginal revenue curve. As such, exogenous production shocks are intrinsically linked to heterogeneous plant behaviour whose implications depend on the demand manifold. As in the benchmark APG decomposition, the demand parameter is extracted from the demand manifold and used to scale $lnTFP^{rr}$ and revenue-function elasticities under curvature restrictions governed by lower (L) and upper (U) bounds as $\gamma_{Hi} \in [L, U] = [2.08, 9.75]$.

Table 4 reports productivity and input reallocation aggregates under varying HARA demand convexity. That is; given the bounds for $\gamma_{Hi}$, the range of average percentage points for technical inefficiency and multifactor reallocation are $(lnTFPQ, lnTFR) \in [-11.16, -11.49] \times [8.69, 9.15]$. The quantitative drivers of APG are insignificantly different from related benchmark values. Intuitively, technical inefficiency and multifactor reallocation are invariant to changes in the curvature under HARA market preferences.

**B. Variation in the Convexity of the Demand Function**

There is also another dimension that involves responsiveness of measures of technical efficiency and input reallocation in relation to firm-size variation that needs exploration. That is, does firm-scale variation in the {Superconvexity, Subconvexity} space of the demand structure produce cross-sectional moments that are at variance with the main results, *ceteris paribus*? For instance, suppose we fix $Q^*_{Hi}$ at the inflection point of Panel D in Figure 6, which reflects the benchmark CES demand function. First, let $Q_{Hi} \neq Q^*_{Hi}$. If $Q_{Hi} < Q^*_{Hi}$, then production occurs in the *superconvex* region of the demand schedule faced by smaller firms, where output size also defines firm-size. However, if $Q_{Hi} > Q^*_{Hi}$, then larger firms face a *subconvex* demand system in the lower region of the demand manifold. For this, we condition on the shape parameter of the demand function taking the mid-point of $\gamma_{Hi} = 5.92$ for the HARA market regime. Setting the data-driven inflection point to $Q^*_{Hi}$=5.84 and adjusting the structure of the demand system enables computation of APG decomposition for both firm-size categories in the HARA economy as shown Table 5.

**Table 4: Sensitivity of Productivity and Reallocation under HARA Characteristics with $\gamma_{Hi} \in [L, U]$ for 49 Four-Digit ISIC Industries**

| PANEL A: Heterogeneous Markups; $\rho_{Hi} = \frac{Q_{Hi}-\gamma_{Hi}\mu_{Hi}}{Q_{Hi}-\gamma_{Hi}\mu_i+Q_{Hi}\mu_{Hi}}$ and $\gamma_{Hi} = 2.08$ | | | | | |
|---|---|---|---|---|---|
| Year | Technical Efficiency | Multifactor Reallocation | Total Labour Reallocation | Paid Employee Reallocation | Capital Reallocation |
| 1995 | -11.19 | 2.53 | 3.71 | 6.61 | -1.18 |
| 1996 | 2.63 | -6.93 | -1.74 | -0.77 | -5.18 |
| 1997 | -4.72 | 19.54 | 23.39 | 23.39 | -3.85 |
| 1998 | 0.73 | -0.84 | 2.75 | 2.75 | -3.59 |
| 1999 | 3.13 | 12.99 | 17.87 | 17.87 | -4.87 |
| 2000 | -88.11 | 51.25 | 0.72 | 0.69 | 50.52 |
| 2001 | 40.65 | 2.66 | -0.73 | -1.64 | 3.40 |
| 2002 | -34.60 | -10.50 | -4.44 | 7.92 | -6.05 |
| 2003 | -20.12 | 7.52 | 4.62 | 1.52 | 2.89 |
| **Mean** | **-11.16** | **8.69** | **4.61** | **5.83** | **3.56** |
| **Median** | **-2.36** | **2.66** | **1.74** | **2.13** | **-3.59** |
| **Std Dev** | **33.23** | **18.46** | **8.94** | **8.47** | **17.94** |
| **PANEL B: Heterogeneous Markups; $\rho_{Hi} = \frac{Q_{Hi}-\gamma_{Hi}\mu_{Hi}}{Q_{Hi}-\gamma_{Hi}\mu_i+Q_{Hi}\mu_{Hi}}$ and $\gamma_{Hi} = 9.75$** | | | | | |
| Year | Technical Efficiency | Multifactor Reallocation | Total Labour Reallocation | Paid Employee Reallocation | Capital Reallocation |
| 1995 | -12.11 | 2.94 | 3.93 | 6.92 | -0.99 |
| 1996 | 2.74 | -7.04 | -1.99 | -0.86 | -5.04 |
| 1997 | -4.89 | 19.71 | 23.40 | 23.40 | -3.68 |
| 1998 | 0.75 | -0.86 | 2.52 | 2.52 | -3.39 |
| 1999 | 3.16 | 12.89 | 20.81 | 20.81 | -7.91 |
| 2000 | -91.18 | 54.48 | 0.34 | 0.35 | 54.14 |
| 2001 | 42.43 | 1.25 | -0.87 | -1.80 | 2.13 |
| 2002 | -35.81 | -9.39 | -5.11 | 8.51 | -4.28 |
| 2003 | -20.01 | 8.45 | 4.79 | 1.51 | 3.65 |
| **Mean** | **-11.49** | **9.15** | **4.78** | **6.13** | **3.84** |
| **Median** | **-2.44** | **2.94** | **1.43** | **2.01** | **-3.39** |
| **Std Dev** | **34.42** | **19.33** | **9.59** | **9.05** | **19.19** |

*Notes*: Producers are price-setters with product differentiation and markup heterogeneity. Scaling revenue function residual and elasticities by the demand parameter equates them to technical efficiency and physical output elasticities, respectively. Numbers in cells are percentage points. The plant-level multifactor productivity uses production function parameters that vary across the 2-digit ISIC code obtained by using Wooldridge (2009); i.e., ivreg29. APG represents aggregate productivity growth rate with entry and exit, and is the aggregate change in final demand *minus* aggregate change in expenditure in inputs, holding inputs constant. Value-added output shares (Domar) are weights. APG decomposes into: technical efficiency and reallocation using the Estimating Equation, Eq. 5 in the text.

**Source**: Author's Calculations.

A characteristic that stands out for smaller firms in the superconvex region of the demand manifold is that when $\gamma_{Hi}$ returns to its midpoint level, all measures of factor-input reallocation are significantly attenuated relative to benchmark levels. At the same time, larger plants in the subconvex demand region experienced a decline in the elasticity of demand with an increase in industry output due to competitive pressures of trade liberalization. The APG contributions of larger plants remained fundamentally invariant to firm-size variation in one sense and diminished in another sense. For instance; while technical inefficiency remains robust to firm size variation, the sensitivity of multifactor input reallocation is scale-dependent as shown by $(lnTFPQ, lnTFR) \in [-12.54,\ -11.16] \times [3.91, 5.56]$. That is, multifactor input reallocation exhibits significant attenuation effects with respect to firms in superconvex and subconvex markets. This pattern of input reallocation behaviour owes its character to the differential speed of factor-input reshuffling across plants. For instance, there is weak reallocation of paid employees for firms facing superconvex demand and weak capital reallocation for firms facing subconvex demand. Both effects had a negative impact on overall factor-input reallocation thereby significantly weakening the robustness of the multifactor reallocation results.

**Table 5: Sensitivity of Productivity and Reallocation under HARA Characteristics by Demand Convexity for 49 Four-Digit ISIC Industries**

| **PANEL A: Heterogeneous Markups for Superconvex Firms; $\rho_{Hi} = \frac{Q_{Hi}-\gamma_{Hi}\mu_{Hi}}{Q_{Hi}-\gamma_{Hi}\mu_i+Q_{Hi}\mu_{Hi}}$ and $\gamma_{Hi} = 5.92$** | | | | | |
|---|---|---|---|---|---|
| **Year** | **Technical Efficiency** | **Multifactor Reallocation** | **Total Labour Reallocation** | **Paid Employee Reallocation** | **Capital Reallocation** |
| 1995 | -14.06 | -0.64 | 0.84 | 0.84 | -1.49 |
| 1996 | 3.14 | -5.56 | -1.28 | -0.24 | -4.27 |
| 1997 | -5.32 | -1.80 | -1.25 | -1.25 | -0.55 |
| 1998 | 0.82 | -3.73 | -1.47 | -1.47 | -2.25 |
| 1999 | 0.84 | 66.11 | 17.62 | 17.62 | 48.48 |
| 2000 | -96.91 | -18.44 | -3.33 | -3.11 | -15.11 |
| 2001 | 48.40 | 11.69 | -2.31 | -2.62 | 14.01 |
| 2002 | -37.79 | -2.83 | -2.42 | -0.03 | -0.41 |
| 2003 | -24.57 | -5.68 | 2.71 | 0.13 | -8.39 |
| **Mean** | **-12.54** | **3.91** | **0.91** | **0.98** | **3.00** |
| **Median** | **-2.66** | **-2.32** | **-1.27** | **-0.13** | **-1.02** |
| **Std Dev** | **37.16** | **23.05** | **6.13** | **5.98** | **17.58** |
| **PANEL B: Heterogeneous Markups for Subconvex Firms; $\rho_{Hi} = \frac{Q_{Hi}-\gamma_{Hi}\mu_{Hi}}{Q_{Hi}-\gamma_{Hi}\mu_i+Q_{Hi}\mu_{Hi}}$ and $\gamma_{Hi} = 5.92$** | | | | | |
| **Year** | **Technical Efficiency** | **Multifactor Reallocation** | **Total Labour Reallocation** | **Paid Employee Reallocation** | **Capital Reallocation** |
| 1995 | -10.87 | -0.42 | 2.97 | 5.91 | -3.39 |
| 1996 | 2.56 | 6.25 | -0.57 | -0.57 | 6.83 |
| 1997 | -4.69 | 23.02 | 24.66 | 24.66 | -1.63 |
| 1998 | 0.72 | 22.99 | 4.12 | 4.12 | 18.86 |
| 1999 | 3.09 | 4.04 | 1.57 | 1.57 | 2.47 |
| 2000 | -87.81 | 3.33 | 3.89 | 3.65 | -0.55 |
| 2001 | 39.71 | 1.91 | 1.51 | 0.91 | 0.40 |
| 2002 | -34.53 | -6.00 | -2.32 | 8.24 | -3.67 |
| 2003 | -19.79 | 0.51 | 2.00 | 1.38 | -1.49 |
| **Mean** | **-11.16** | **5.56** | **3.78** | **4.99** | **1.78** |
| **Median** | **-2.34** | **2.62** | **1.78** | **2.61** | **-0.27** |
| **Std Dev** | **32.96** | **9.75** | **7.60** | **7.43** | **6.72** |

*Notes*: Producers are price-setters with product differentiation and markup heterogeneity. Scaling revenue function residual and elasticities by the demand parameter equates them to technical efficiency and physical output elasticities, respectively. Numbers in cells are percentage points. The plant-level multifactor productivity uses production function parameters that vary across the 2-digit ISIC code obtained by using Wooldridge (2009); i.e., ivreg29. APG represents aggregate productivity growth rate with entry and exit, and is the aggregate change in final demand *minus* aggregate change in expenditure in inputs, holding inputs constant. Value-added output shares (Domar) are weights. APG decomposes into: technical efficiency and reallocation using the Estimating Equation, Eq. 5 in the text.

**Source**: Author's Calculations

However, an unexpected result emerges from the comparison of labour $(lnLR)$ and capital $(lnKR)$ reallocation under CES with the subconvex demand benchmark. To clearly capture this connection, consider the relationship $lnLR^{ces} \times lnKR^{ces} = 2.90 \times 2.14$ versus $LR^{sub} \times lnKR^{sub} = 3.78 \times 1.78$.

The latter comparator group; of course, consists larger firms. A closer look at the relevant markets shows a close approximation of factor reallocation between all firms facing CES demand and those facing subconvex demand; i.e., $lnLR^{ces} \cong lnLR^{sub}$ and $lnKR^{ces} \cong lnKR^{sub}$ . Thus, factor-input reallocation in the CES market is robust to input-reallocation associated with larger firms in subconvex markets. One explanation for the proximate orders of magnitude in resource reshuffling across market conventions is that larger plants in the subconvex regime charge predominantly constant markups. On the other hand, smaller firms facing the superconvex demand system operated largely under autarky in a period of trade reforms and experienced acutely subdued labour reallocation while enjoying robust capital reallocation at $LR^{sup} \times lnKR^{sup} = 0.91 \times 3.00$.

### C. Plant Expansion ($\Delta \boldsymbol{E} > \mathbf{0}$) versus Plant Contraction ($\Delta \boldsymbol{E} \leq \mathbf{0}$)

Holding constant the demand curvature at its midpoint level, we separate downsizing plants from expanding ones to further assess the robustness of the benchmark results. First and foremost, first-order moments for technical efficiency are fundamentally negative. We therefore consider two characterizations of firms: downsizers $(\Delta E \leq 0)$ and expanding firms $(\Delta E > 0)$. Table 6 reports first-order moments of multifactor reallocation results for shrinking and growing plants over the sample period as $lnTFR \in [-0.33, 12.07]$, respectively. Notably, related technical efficiency $lnTFPQ \in [-12.20,\ -11.29]$ is robust to the benchmark level of -11.41%. Because technical efficiency is about producing more with less, the combination of technical efficiency deterioration and downsizing of plants (also referred to as unsuccessful downsizers) has at least three potential explanations. First, unsuccessful downsizers may have experienced a decline in the demand for their products and increasing returns to scale. Second, the failure to improve production efficiency through introduction of leaner and meaner establishments was associated with elastic demand for product varieties. Third, firms faced deteriorating demand for products and experienced incomplete labour input adjustment; a plausible supposition for most industries in Eswatini, we would argue. Moreover, unsuccessful downsizers moved paid employment resources to worse performing plants in the sample.

A full characterization of the growing class of firms is that paid labour reallocation was consistently positive every year for upsizing plants with an average very close to benchmark levels; i.e., $lnLR = 8.29$. Again, one potential explanation could be that: First, firms experienced long-term negative production efficiency shocks and inelastic demand for their product varieties. Second, there was growing demand in the face of diminishing returns to scale. Third, trade liberalization may have induced multinationals to relocate to the neighbouring larger market with their skilled workers while remaining firms transitioned to lower quality labour inputs; a reasonable supposition informed by anecdotal evidence on local industries.

## D. HARA Demand Elasticities and Markups

**Table 6: Sensitivity of Productivity and Reallocation under HARA Characteristics with Labour Adjustments for 49 Four-Digit ISIC Industries**

| **PANEL A: Heterogeneous Markups with $\Delta E \leq 0$; $\rho_{Hi} = \frac{Q_{Hi}-\gamma_{Hi}\mu_{Hi}}{Q_{Hi}-\gamma_{Hi}\mu_i+Q_{Hi}\mu_{Hi}}$ and $\gamma_{Hi} = 5.92$** | | | | | |
|---|---|---|---|---|---|
| **Year** | **Technical Efficiency** | **Multifactor Reallocation** | **Total Labour Reallocation** | **Paid Employee Reallocation** | **Capital Reallocation** |
| 1995 | -11.40 | -1.76 | -1.74 | -1.74 | -0.01 |
| 1996 | 3.18 | -3.44 | -6.82 | -6.82 | 3.38 |
| 1997 | -4.64 | -1.82 | -1.89 | -1.89 | 0.07 |
| 1998 | 0.84 | -3.28 | -2.44 | -2.44 | -0.83 |
| 1999 | 2.95 | 0.36 | -0.69 | -0.69 | 1.06 |
| 2000 | -88.49 | 1.10 | -4.11 | -4.11 | 5.21 |
| 2001 | 40.11 | 11.17 | -2.73 | -3.53 | 13.91 |
| 2002 | -30.37 | -3.15 | -2.96 | -0.92 | -0.19 |
| 2003 | -22.04 | -2.20 | -2.27 | -1.01 | 0.06 |
| **Mean** | **-12.20** | **-0.33** | **-2.85** | **-2.57** | **2.51** |
| **Median** | **-4.64** | **-1.82** | **-2.44** | **-1.89** | **0.07** |
| **Std Dev** | **34.78** | **4.59** | **1.75** | **1.97** | **4.70** |
| **PANEL B: Heterogeneous Markups with $\Delta E > 0$; $\rho_{Hi} = \frac{Q_{Hi}-\gamma_{Hi}\mu_{Hi}}{Q_{Hi}-\gamma_{Hi}\mu_i+Q_{Hi}\mu_{Hi}}$ and $\gamma_{Hi} = 5.92$** | | | | | |
| **Year** | **Technical Efficiency** | **Multifactor Reallocation** | **Total Labour Reallocation** | **Paid Employee Reallocation** | **Capital Reallocation** |
| 1995 | -11.47 | 1.05 | 5.56 | 8.50 | -4.50 |
| 1996 | 2.60 | 4.25 | 4.96 | 6.00 | -0.70 |
| 1997 | -4.72 | 23.11 | 25.30 | 25.30 | -2.19 |
| 1998 | 0.71 | 22.29 | 5.10 | 5.10 | 17.18 |
| 1999 | 2.11 | 69.06 | 19.89 | 19.89 | 49.16 |
| 2000 | -84.48 | 2.31 | 4.88 | 4.65 | -2.56 |
| 2001 | 41.27 | 2.07 | 1.81 | 1.81 | 0.26 |
| 2002 | -38.79 | -5.50 | -0.50 | 9.13 | -5.00 |
| 2003 | -20.15 | 2.08 | 3.75 | 2.53 | -1.67 |
| **Mean** | **-11.29** | **12.07** | **7.07** | **8.29** | **4.99** |
| **Median** | **-2.36** | **2.19** | **4.92** | **5.55** | **-1.18** |
| **Std Dev** | **32.73** | **22.12** | **8.55** | **8.14** | **16.74** |

*Notes*: Producers are price-setters with product differentiation and markup heterogeneity. Scaling revenue function residual and elasticities by the demand parameter equates them to technical efficiency and physical output elasticities, respectively. Numbers in cells are percentage points. The plant-level multifactor productivity uses production function parameters that vary across the 2-digit ISIC code obtained by using Wooldridge (2009); i.e., ivreg29. APG represents aggregate productivity growth rate with entry and exit, and is the aggregate change in final demand *minus* aggregate change in expenditure in inputs, holding inputs constant. Value-added output shares (Domar) are weights. APG decomposes into: technical efficiency and reallocation using the Estimating Equation, Eq. 5 in the text.

**Source**: Author's Calculations.

The measurement of demand elasticities and markups for products is important for understanding the behaviour of firms in superconvex and subconvex markets as well as the implicit patterns of consumer product substitutability. Table 7 reports HARA median demand elasticities and associated median markups for variously arranged classes of producers. The results exhibit demand inelasticities and markups for both inefficient and efficient firms as $(\varepsilon_{DHi}, 1 + \mu_{Hi}) \in [-0.65, -0.69] \times [2.84, 3.20]$, respectively. These demand inelasticities imply low product substitutability across product varieties.[22] Low-substitutability industries are necessarily associated with high dispersion in technical efficiency and low average efficiency levels in production (Syverson (2004), a pattern that fits the description of Eswatini. Consumers therefore do not easily switch between product varieties, and producers are able to charge higher markups. In line with the literature, there are significant markup differences between unproductive and productive firms. Firms with negative technical efficiency price their product varieties less than firms with positive technical efficiency. These patterns are consistent with firms investing in capital and human capabilities to raise their productivity and subsequently charge higher markups.[23] A crucial comparative static is that firms facing subconvex demand exhibit elasticity of demand that is decreasing in industry output.

**Table 7: Demand Elasticities and Markups by Productivity, Employment, and Production**

| **Classification** | **Median Elasticities of Demand ($\varepsilon_{DHi}$)** | | **Median Markups ($1+\mu_i$)** | |
|---|---|---|---|---|
| | $lnTFPQ \leq 0$ | $lnTFPQ > 0$ | $lnTFPQ \leq 0$ | $lnTFPQ > 0$ |
| All Firms | -0.68 | -0.69 | 3.11 | 3.16 |
| $\Delta E \leq 0$ | -0.65 | -0.69 | 2.84 | 3.20 |
| $\Delta E > 0$ | -0.67 | -0.69 | 3.00 | 3.11 |
| $Q \leq Q^*$ | -0.76 | -0.76 | 3.87 | 3.70 |
| $Q > Q^*$ | -0.67 | -0.68 | 2.99 | 3.05 |

Note: The inflection point for industry output is common at $Q^* = 5.84$ where $\varepsilon_{DHi} = \rho_{Hi} - 1$ and $1 + \mu_i$ are defined in Table 2 for firms facing HARA demand technologies. $\Delta E$ denotes a change in the employment of Paid Labour. The midpoint measure of curvature of the demand function is maintained.

**Source**: Author's Calculation.

In the case of smaller downtown 'boutique' plants facing inelastic superconvex demand, their inelasticities and markups are $(\varepsilon_{DHi}, 1 + \mu_{Hi}) \in [-0.76] \times [3.87, 3.70]$. The small inefficient plants (4th row) charge even higher markups of 3.87 relative to the more technically efficient counterparts at 3.70. These firms have relatively weaker demand inelasticity but higher markups. More specifically, a 10 percent increase in the price of product varieties by these firms reduces the quantity demanded by only 7.6 percent. At the same time, the small unproductive plants charge 17 percent more in markups than their efficient counterparts. Though it sounds paradoxical that inefficient firms of whatever size should price their products higher than efficient plants of equivalent size, this result has *a priori* foundations in Mrázová and Neary (2017) and Dhingra and Morrow (2019). Under regularity conditions, smaller plants do charge higher markups. An

[22] In Syverson (2004), product substitutability is negatively correlated with within-industry productivity dispersion and positively correlated with median productivity in a study of 443 US manufacturing industries.

[23] As partly explained in the proof-of-concept discussion, Foster et al. (2018) reports a pooled markup of 2.08 corresponding to the demand elasticity of -1.93, and a common markup of 1.52 in the concrete industry corresponding to the demand elasticity of -2.92. Aghion et al. (2008) concludes that South African markups in the various manufacturing industries as higher than markups in corresponding industries world-wide. Moreover, manufactured products in Eswatini are subject to autarkic prices and prices that obtain in SACU, preferential foreign and free world markets, the sectoral aggregates of which therefore likely resemble the South African patterns. In 1987, Oberfield and Raval (2021) found that demand elasticities ranged between 3 and 7 and fell to 3.3 in 2007 for the US manufacturing sector.

important comparative static for this class of firms is that superconvexity in demand implies that the elasticity of demand increases in industry output.

An interesting result concerns subconvex markets with demand inelasticities and high markups (5th row) that mimic those of the entire productive and unproductive survey of firms (1st row). The orders of markup magnitude within and between technical efficiency classes move in tandem. One explanation for this property is that idiosyncratic shocks to large plants propagate to aggregate markup fluctuations within various industries. Perhaps this is validation of the granularity proposition by Gabaix (2011).

The next subsection relies on economic intuition to identify and measure the effects of potential sources of misallocation in factor-inputs for resources-constrained and unconstrained incumbent firms.

### 5.3. Effects of Misallocation Sources on Input Distortions

The APG decomposition results reveal that input reallocation remains an important channel among incumbent firms. In what follows is a dissection of the dataset into input-constrained and input-unconstrained firms to recover the relationship between sources of misallocation and factor-input wedges. Although there is a range of candidate misallocation sources, our analysis isolates firm-level pricing power, technical efficiency, and factor intensity for this investigation.[24] Producer market power is included because of its systematic variability that causes market-shares to reallocate from low- to high-markup firms (De Loecker et al., 2020; and Peters, 2020). Plant heterogeneity in productivity facilitates growth and survival of high technical efficiency plants while unproductive establishments contract and exit the market (Nishida et al., 2014; and Ho et al., 2019). In a well-functioning market economy, these Darwinian forces lead to increases in APG. Factor intensity reflects the extent of input misallocation and gives insight on APG contribution coming from specific factor-inputs (Ho et al., 2019; and Oberfield and Raval, 2021). We therefore adopt and estimate the following fixed effects model

$$\tau_{it}^{m} = \beta_{year} + I\{\tau_{it}^{m} \geq 0\} \times \left\{ \beta_{lnTFPQ}^{+} lnTFPQ_{it-1} + \beta_{K/L}^{+} \, ln\left( \frac{K_{it-1}^{PME}/PE}{L_{it-1}} \right) + \beta_{\mu}^{+}(1+\mu_{it-1}) \right\}$$

$$+ I\{\tau_{it}^{m} < 0\} \times \left\{ \beta_{lnTFPQ}^{-} lnTFPQ_{it-1} + \beta_{K/L}^{-} \, ln\left( \frac{K_{it-1}^{PME}/PE}{L_{it-1}} \right) + \beta_{\mu}^{-}(1+\mu_{it-1}) \right\} + \omega_{it} \qquad [8]$$

where $\beta_{year}$ denotes year fixed-effects and $\omega_{it}$ is random noise. The indicator function $I(\cdot)$ is equal to one if the argument holds and zero otherwise. The interaction terms distinguish between I nput-constrained and input-unconstrained regimes of plants together with associated input wedges.

[24] Other misallocation sources include financial credit frictions (Buera, et al., 2011; and Gopinath et al., 2017), as well as adjustment costs and informational frictions (David and Venkateswaran, 2019).

A closer look at the distribution of undersized plants, or $\tau_{it}^{m} \geq 0$, (not reported due to space constraints) shows that input-constrained firms are a majority relative to oversized plants, or $\tau_{it}^{m} < 0$. Table 8 reports the regression results.

**Input-Constrained Plants**, **or** $\tau_{it}^{m} \geq 0$: The estimated distortionary effects of capital intensity, or a positive $\beta_{K/L}^{+}$, means high capital growth firms are constrained by inefficient allocation of factor-inputs. As evidence of *relative* labour market flexibility, a one-percentage point increase in capital intensity increases labour distortions by 4.95 percentage points and reduces capital distortions by 45.59 percentage points for the under-sized plants. That is, resource-constrained but capital-intensive plants have higher marginal products of labour and some leverage for labour adjustments. These plants also experienced higher labour distortions. Notwithstanding their size-limitation, plants are subject to significantly lower capital distortions, reflecting the irreversibility constraints of capital investments that characterize developing economies (Ho et al., 2019).

**Table 8: Regression Analysis of Input Distortions on Misallocation Sources**

| Variables | **Labour Distortions** | **Capital Distortions** |
|---|---|---|
| | b/se | b/se |
| $lnTFPQ_{it-1} \mid \tau_{it}^{m} \geq 0$ | -1.781 | -4.458 |
| | (1.7596) | (10.8692) |
| $ln\left(\overset{PME}{K_{it-1}}/\overset{PE}{L_{it-1}}\right) \mid \tau_{it}^{m} \geq 0$ | 4.947*** | -45.587*** |
| | (1.0382) | (7.7444) |
| $(1+\mu_{it-1}) \mid \tau_{it}^{m} \geq 0$ | 1.403 | -2.426 |
| | (3.1788) | (14.9061) |
| $lnTFPQ_{it-1} \mid \tau_{it}^{m} < 0$ | -1.279 | 3.217 |
| | (2.2084) | (9.5605) |
| $ln\left(\overset{PME}{K_{it-1}}/\overset{PE}{L_{it-1}}\right) \mid \tau_{it}^{m} < 0$ | 0.284 | -36.842*** |
| | (1.1188) | (6.2221) |
| $1+\mu_{st-1} \mid \tau_{it}^{m} < 0$ | -2.087 | 1.538 |
| | (2.8073) | (15.1794) |
| Year Fixed-Effects | Yes | Yes |

Note: *p****<0.001.

**Source**: Author's calculations.

One explanation of these findings is that the interaction of factor input markets through substitutability generated an asymmetric input mix that exacerbated factor-input distortions. As Oberfield and Raval (2021) observe, the elasticity of substitution between capital and labour the extent of which is explained by heterogeneity in capital intensities represents substitution within firms and reallocation across firms. Furthermore, as argued by Edwards et al. (2013), the manufacturing sector experienced an uncompetitive investment environment, restrictive and distortionary regulations, unsuitable state intervention inhibited the emergence of private firms and the expansion of existing industries. This partly explains the dominance of undersized plants that are heterogeneous in capital intensities.

**Plants with unconstrained Inputs, or $\tau_{it}^{m} < 0$**: Turning to the capital-labour ratio of plants with unconstrained inputs, higher $\beta_{K/L}^{-}$ indicates that oversized firms have less surplus capital and more surplus labour. This is consistent with the economic notion that low $lnTFPQ$ firms are associated with low marginal products of inputs and therefore are unable to leverage on factor-inputs to achieve scale economies. Therefore, the extent of misallocation declines without variability in $lnTFPQ$ for undersized establishments. In the case of oversized plants, factor input reallocation declines without much variation in technical efficiency. The capital intensity coefficient on capital distortions is significantly negative relative to the labour counterpart, signifying the capital flight experienced during the period of trade liberalization in the Customs Union. If we interpret PME capital investment as a fixed/sunk cost, then a one-percentage increase in the FSC of capital relative to labour reduced capital distortion by 36.84 percentage points. The intuition for this is that the non-robust capital inflows of the 1990s and 2000s precipitated labour substitution for capital.

## 6 Summary and Conclusion

This article studied the decomposition of structural aggregate productivity growth (APG) under differences market conditions with demand- and supply-side controls, determined comparative statics predictions for firms and economic outcomes, and examined patterns of input distortions. There is enormous value in studying APG decomposition under monopolistic competition to determine the effects of price-setting on within-firm productivity and factor-input reallocation. A second-order benefit of micro price-setting schemes is the ability to back out demand manifolds usable as sufficient statistics for comparative statics predictions covering a variety of pertinent issues.

This paper finds technical efficiency decline of -11.41% and productivity-enhancing resource reallocation of 8.91% by moving from a price-taking environment to a HARA markup pricing market. The results on technical inefficiency are robust to demand curvature dynamics, demand function convexity and plant-size variation. Factor-input reallocation results are insensitive to variations in demand curvature and plant-size dynamics, but sensitive to demand function convexity and firm contraction. Moreover, downsizers experienced severe input misallocation from productive to unproductive business units within industries. Input reallocation under CES preferences remained robust to input reallocation among larger firms facing subconvex demand. Smaller firms operated largely under autarky in a market size of approximately One Million people during a period of trade reforms and experienced acutely subdued labour reallocation while exhibiting robust capital reallocation.

Looking at comparative statics predictions, a key result is that plants with attenuated weaknesses in productivity charge higher markups in high capital-intensity industries. Similarly, the elasticity of demand increases (decreases) with industry output for smaller (larger) firms. Consistent with Marshal's Second Law of Demand, the demand elasticity also increases with markups and, by implication, increases with prices along the demand curve. Finally, an increase in capital-intensity correlates with varied adjustments in factor-input distortions for resource-constrained and unconstrained plants, *ceteris paribus*.

The lacklustre APG performance is consistent with the current extent of export product diversification presented in the background section of this article. The underlying message from these results points to the *prima facie* need for urgent identification of new products that are close to the globally connected core, attract firms of suitable size with multiproduct capacity to produce the identified products and participate in global value chains (see, for example, Herrendorf et al., 2013; and Herrendorf et al., 2015). The economy needs to focus on building *relevant* productive capability in technologies, skilled manpower, capital stock, and quality institutions (Hausmann and Hidalgo, 2010; 2011). Perhaps the country's appeal to potential entry of multinational enterprises rests on having a political environment and macroeconomic stability that are conducive to business enterprise. At industrial level, the combined effects of input misallocation, technical efficiency losses, and weak factor-input reallocation call for improvement in the flexibility of factor markets as well as withdrawal of distortionary state participation in business enterprise (Edwards et al., 2013).

**REFERENCES**

Ackerberg D. A., K. Caves, and G. Frazer (2015). Identification Properties of Recent Production Function Estimators. *Econometrica*, 83 (6), 2411-2451. https://doi.org/10.3982/ECTA13408.
Aghion, P., M. Braun, and J. Fedderke (2008). Competition and Productivity Growth in South Africa. *Economics of Transition* 16(4), 741-768. http://dx.doi.org/10.1111/j.1468-0351.2008.00336.x.
Asplund, M. and V. Nocke (2006). Firm Turnover in Imperfectly Competitive Markets. *Review of Economic Studies* **73**, 295-327. https://doi.org/10.2139/ssrn.393880.
Baker J. B. (2007). Market Definition: An Analytical Overview. *Antitrust Law Journal*, 74 (1), 129-173.
Bartelsman E, J. Haltiwanger, and S. Scarpetta (2004) Microeconomic Evidence of Creative Destruction in Industrial and Developing Countries. World Bank policy research working paper 3464.
Bertola G. and R. J. Caballero (1994). Irreversibility and Aggregate Investment. *The Review of Economic Studies*, 61 (2), 223–246, https://doi.org/10.2307/2297979.
Bigsten A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J.W. Gunning, R. Oostendorp, C. Pattillo, M. Soderbom and F. Teal (2005). Adjustment Costs and Irreversibility as Determinants of Investment: Evidence from African Manufacturing. *Contributions to Economic Analysis & Policy*, 4(1), 12, 1-27. https://doi.org/10.2202/1538-0645.1228.
Bils M., P. J. Klenow, and C. Ruane (2021). Misallocation or Mismeasurement? *Journal of Monetary Economics*, 124S, S39–S56, https://doi.org/10.1016/j.jmoneco.2021.09.004.
Buera F. J., J. P. Kaboski, and Y. Shin (2011). Finance and Development: A Tale of Two Sectors. *The American Economic Review*. 101 (5), 1964-2002, https://doi.org/10.1257/aer.101.5.1964.
Burya A. and M. Mishra (2022). Variable Markups, Demand Elasticity and Pass-through of Marginal Costs into Prices. Mimeograph.
Cadot O., C. Carrefre, and V. Strauss-Kahn (2011). Export Diversification: What's Behind the Hump? *The Review of Economics and Statistics*, 93(2), 590–605. https://doi.org/10.1162/RESTa00078
Costinot A. (2009). An Elementary Theory of Comparative Advantage. *Econometrica*, 77 (4), 1165-1192. http://dx.doi.org/10.3982/ECTA7636.
David, J.M., and V. Venkateswaran, (2019). The Sources of Capital Misallocation. *American Economic Review*, 109 (7), 2531–2567. https://DOI.10.1257/aer.20180336.
De Loecker J., J. Eeckhout, and G. Unger (2020). The Rise of Market Power and The Macroeconomic Implications. *The Quarterly Journal of Economics*, 135 (2), 561–644. https://doi.org/10.1093/qje/qjz041.
De Loecker J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, Markups, and Trade Reform. *Econometrica*, 84 (2), 445–510. https://doi.org/10.3982/ECTA11042
de Vries G., M. Timmer and K. de Vries (2015). Structural Transformation in Africa: Static Gains, Dynamic Losses. *The Journal of Development Studies*, 51 (6), 674–688. http://dx.doi.org/10.1080/00220388.2014.997222.
Demirer M. (2022). Production Function Estimation with Factor-Augmenting Technology: An Application to Markups. MIT Sloan School of Management.
Dhingra S. and J. Morrow (2019). Monopolistic Competition and Optimum Product Diversity under Firm Heterogeneity. *Journal of Political Economy,* 127 (1): 196-232. https://doi.org/10.1086/700732
Edwards L., F. Flatters, M. Stern and Y. Ramskolowan (2013). Swaziland Economic Diversification Study: Final Report", *DNA Economics.*
Ericson, R. and Pakes A. (1995). Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies,* 62(1), 53-82. http://dx.doi.org/10.2307/2297841
Forster, L., Haltiwanger, J. and Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *The American Economic Review*, 98(1), 394-425. http://dx.doi.org/10.1257/aer.98.1.394
Foster L. S., C. A. Grim, J. Haltiwanger and Z. Wolf (2017). Macro and Micro Dynamics of Productivity: From Devilish Details to Insights, NBER Working Paper Series.
Foster, L., Grim, C., Haltiwanger, J. and Wolf, Z. (2016). Firm-Level Dispersion in Productivity: Is the Devil in the Details? *The American Economic Review*, 106(5), 95-98. https://doi.org/10.1257/aer.p20161023
Foster, L., R. Kulick and C. Syverson (2018). Misallocation Measures: The Distortion That Ate the Residual. NBER Working Paper 24199.
Fozzi A. and F. Schivardi (2016). Demand or productivity: what determines firm growth? RAND Journal of Economics, 47 (3), 608–630. https://doi.org/10.1111/1756-2171.12142.

Gabaix X (2011) The Granular Origins of Aggregate Fluctuations. *Econometrica*, 79(3), 733–772. https://doi.org/10.3982/ECTA8769

Gandhi A., S. Navarro and D. A. Rivers (2020). On the Identification of Gross Output Production Functions, Journal of Political Economy, 128 (8). http://dx.doi.org/10.1086/707736

Gopinath G., S. Ebnem Kalemli-¨Ozcan, L. Karabarbounis, And C. Villegas-Sanchez (2017). Capital Allocation and Productivity in South Europe. *The Quarterly Journal of Economics:* 1915–1967. https://doi.10.1093/qje/qjx024.

Hall R. E. and C. I. Jones (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics*, 114 (1), 83-116. https://doi.org/10.1162/003355399555954

Haltiwanger J., Kulick R., and Syverson C. (2018). Misallocation Measures: The Distortion That Ate the Residual, NBER, Cambridge.

Hartmann D., M. R. Guevara, C. Jara-Figueroa, M. Aristaran and C. A. Hidalgo (2017). Linking Economic Complexity, Institutions, and Income Inequality. *World Development*, 93, 75–93. http://dx.doi.org/10.1016/j.worlddev.2016.12.020.

Hausmann R. (2016). Economic Development and the Accumulation of Know-how. *Welsh Economic Review*, 24, 13-16.

Hausmann R. and C. A. Hidalgo (2010). Country Diversification, Product Ubiquity, and Economic Divergence. CID Working Paper No. 201.

Hausmann R. and C. A. Hidalgo (2011). The network structure of economic output. *Journal of Economic Growth*, 16 (4), 309-342. https://doi.org/10.1007/s10887-011-9071-4.

Herrendorf B., R. Rogerson, and Á. Valentinyi (2013). Two Perspectives on Preferences and Structural Transformation, *American Economic Review,* 103 (7): 2752-2789. https://dx.doi.org/10.1257/aer.103.7.2752.

Herrendorf B., R. Rogerson, and Á. Valentinyi (2015). Sectoral Technology and Structural Transformation, *American Economic Journal: Macroeconomics*, Vol. 7(4), 104-133. http://dx.doi.org/10.1257/mac.20130041.

Ho A. T. Y., K. P. Huynh, and D. T. Jacho-Chávez (2019). Productivity and Reallocation: Evidence From Ecuadorian Firm-Level Data. *Economia*, 83-110. https://doi.org/10.1353/eco.2019.0009.

Hopenhayn, H. A. (1992). Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica*, 60(5): 1127-1150. https://doi.org/10.2307/2951541
https://doi.org/10.1007/s10290-017-0303-3.

Hulten C. R. (1978). Growth Accounting with Intermediate Inputs, *Review of Economic Studies*, 45: 511–518. https://doi.org/10.2307/2297252.

Jovanovic, B (1982). Selection and Evolution of Industry. *Econometrica*, 15 (3), 649-670. https://doi.org/10.2307/1912606

Kiviet J. F. (2020). Microeconometric Dynamic Panel Data Methods: Model Specification and Selection Issues. *Econometrics and Statistics*, 13, 16–45. https://doi.org/10.1016/j.ecosta.2019.08.003.

Klette T. J. and Z. Griliches (1996). The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous. *Journal of Applied Econometrics*, 11(4), -343 - 361. https://doi.org/10.1002/(SICI)1099-1255(199607)11:4%3C343::AID-JAE404%3E3.0.CO;2-4

Kumar P. and H. Zhang (2019). Productivity or Unexpected Demand Shocks: What Determines Firms' Investment and Exit Decisions? *International Economic Review*, 60 (1). https://doi.org/10.1111/iere.12354.

Kwon H.U., F. Narita, and M. Narita (2015). Resource Reallocation and Zombie Lending in Japan in the 1990s. *Review of Economic Dynamics,* 8 (4), 709-732. https://doi.org/10.1016/j.red.2015.07.001.

Levinsohn, J. and A. Petrin (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies,* 7: 317-341. https://doi.org/10.1111/1467-937X.00246.

Lucas R. Jr (1977). Understanding business cycles, *Carnegie-Rochester Conference Series on Public Policy*. https://doi.org/10.1016/0167-2231(77)90002-1.

Matsuyama K. and P. Ushchev (2022). Selection and Sorting of Heterogeneous Firms through Competitive Pressures. Northwestern University, USA and ECARES, Université Libre de Bruxelles, Belgium.

McMillan M. and A. Zeufack (2022). Labor Productivity Growth and Industrialization in Africa. *Journal of Economic Perspectives*. 36 (1), 3–32. https://doi.org/10.1257/jep.36.1.3.

McMillan M., D. Rodrik, and I´. Verduzco-Gallo (2014). Globalization, Structural Change, and Productivity Growth, with an Update on Africa, *World Development*, 63: 11–32. http://dx.doi.org/10.1016/j.worlddev.2013.10.012.

Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695-1725. https://doi.org/10.1111/1468-0262.00467

Mhlanga S. V. and N. R. Rankin (2021). Fixed Costs, Markups and Concentration in Eswatini (Swaziland): A Firm-Level Analysis of Panel Data. *South African Journal of Economics*, 89, 391-416. https://doi.org/10.1111/saje.12289.

Mrázová M. and J. P. Neary (2017). Not So Demanding: Demand Structure and Firm Behaviour. *American Economic Review,* 107(12): 3835–3874. http://dx.doi.10.1257/aer.20160175.

Mrázová M., J. P. Neary, and M. Parenti (2021). Sales and Markup Dispersion: Theory and Empirics. *Econometrica*, 89 (4), 1753-1788. https://doi.org/10.3982/ECTA17416.

Nishida, M., A. Petrin and S. Polanec (2014). Exploring Reallocation's Apparent Weak Contribution to Growth. *Journal Productivity Analysis,* 42(2): 187-210. https://doi.org/10.1007/s11123-013-0380-9.

Oberfield E. and D. Raval (2021). Micro Data to Macro Technology. *Econometrica*, 89 (2): 703-732. https://doi.org/10.3982/ECTA12807

Olley, S., and A. Pakes (1996). The Dynamics of Productivity in The Telecommunications Equipment Industry. *Econometrica,* 64, 1263– 1298. https://doi.org/10.2307/2171831.

Pagés C., G. Pierre, and S. Scarpetta (2009). Job Creation in Latin America and the Caribbean: Recent Trends and The Policy Challenges. Macmillan, New York.

Pan Q. (2022). Identification of Gross Output Production Functions with a Nonseparable Productivity Shock, Department of Economics, University of Texas.

Pavcnik, N. (2002). Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants. *The Review of Economic Studies*, 69(1), 245-276. https://doi.org/10.1111/1467-937X.00205.

Perets G. S. and E. Yashiv (2015). The Fundamental Nature of HARA Utility. Discussion Papers 1522, Centre for Macroeconomics (CFM).

Peters M. (2020). Heterogeneous Markups, Growth, and Endogenous Misallocation. *Econometrica*, 88 (5), 2037–2073. https://doi.org/10.3982/ECTA15565.

Petrin A., T. K. White, and J. P. Reiter (2011). The Impact of Plant-Level Resource Reallocations and Technical Progress on U.S. Macroeconomic Growth. *Review of Economic Dynamics*, 14: 3–26. https://dx.doi.org/10.1016/j.red.2010.09.004.

Petrin, A. and J. Levinsohn (2012). Measuring Aggregate Productivity Growth Using Plant-Level Data. *RAND Journal of Economics*, 43(4), 705–725. https://doi.org/10.1111/1756-2171.12005.

Pindyck R. S. (1991). Irreversibility, Uncertainty, and Investment. *Journal of Economic Perspectives*, XXIX, 1110-1148. https://doi.org/10.2307/1885568.

Pollak R. A. (1971). Additive Utility Functions and Linear Angel Curves. *The Review of Economic Studies*, 38 (4), 401-414.

Raval, D.R. (2019). The Micro Elasticity of Substitution and Non-Neutral Technology. *The RAND Journal of Economics*, 50(1), 147–167. http://doi.org/10.1111/1756-2171.12265.

Roodman D. (2009a). How to do Xtabond2: An Introduction to Difference and System GMM in Stata. *The Stata Journal*, 9 (1), 86–136, https://doi.org/10.1177%2F1536867X0900900106.

Roodman D. (2009b). PRACTITIONERS' CORNER: A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics*, 71, 135-158. https://doi.org/10.1111/j.1468-0084.2008.00542.x.

United Nations Economic Commission for Africa (2018). Eswatini Structural Transformation, Employment, Production and Society.

Weintraub G.Y., C. L. Benkard, and B. Van Roy (2011). Industry Dynamics: Foundations for Models with an Infinite Number of Firms. *Journal of Economic Theory*, 146, 1965–1994. https://doi.org/10.1016/j.jet.2011.05.007.

Wooldridge, J.M. (2009). On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables. *Economics Letters*, 104(3), 112–114. http://dx.doi.org/10.1016/j.econlet.2009.04.026

Zhelobokdo E., S. Kokovin, M. Parenti, and J-F Thisse (2012). Monopolistic Competition: Beyond the Constant Elasticity of Substitution. *Econometrica*, 80 (6), 2765–2784. http://dx.doi.org/10.2307/23357240.