HOW PRODUCT ATTRIBUTES AFFECT KPI: A CASE STUDY OF A 500-WORKER METAL STAMPING FACTORY IN CHINA

Raymond Cheng,
Pit-yan Li

ABSTRACT

While Key Performance Indicators (KPIs) can best be constructed following Doran's S.M.A.R.T. concept (1981), not every manufacturing business could afford to have the luxury to accurately measure them, especially when these businesses could be SMEs operating with limited resources. Decoene and Bruggeman (2006), therefore, suggested the use of balanced scorecard to help align business strategy with managers' motivation in measuring performance yet such an idea could only work for large enterprises because 'substantial barriers to strategic performance management' do exist for SMEs (Hudson, Smart, & Bourne, 2001) and that 'actionable knowledge' could be hard to acquire when performance for SMEs in itself can already be difficult to measure (Meyer, 2005). This paper, hence, proposed measuring a list of standardized and easy-to-measure KPIs as well as some simple attributes of the products to see if there is any relationship between them. Data were recorded from a relatively small-scale, traditional 500-worker metal stamping factory in China over a period of 9 months (involving 98 shipments of approximately 250,000 pieces of finished products). The results, using multiple t-tests and k-means clustering, were analyzed to see how different combination of attributes of the respective products could affect (and hence change) the resulting KPIs. Product attributes such as size, complexity (and hence allocation of manpower and production time), pricing and order types (e.g. new vs. repeat orders, urgency of order, etc.) were used to compare with the various standard KPIs like defect rate, yield loss, productivity, worker efficiency, etc. The findings revealed, in fact, that when measuring performance alone can already be a complicated issue of people management, psychology, sales and marketing, as well as project management (Spitzer, 2007, p.288), continuous improvement in business performance, product quality, and hence profitability can be realistically achieved through boosting the required KPIs (say, the defect rate) via properly and carefully controlling the combination of product attributes, and hence, the corresponding sales and marketing strategies.

Keywords**:** key performance indicator, KPI, manufacturing, product attribute, defect rate

Overview of the Practical Dilemma

Good production managers are always on the search for better combinations of Key Performance Indicators (KPIs) that can be used to effectively measure the performance and efficiency of the manufacturing team and to accurately reflect and benchmark performance levels against best and worst in-class performers. Selection of KPIs, however, is often, if not mostly, derived or drawn from a much larger pool of other previously tested KPIs that are more commonly known to production managers. And knowing the fact that these metrics should best be aligned toward the organization's core business strategy (Lambert, 2001) and that any move towards increased transparency and accountability via the adoption of any new management metrics or KPIs could be interpreted as threatening[1] within an organization if it has not been previously experienced (Stivers & Joyce, 2000), a clear 'line of sight linkage between the various business levels' (Keegan, Eiler, & Jones, 1989) is often difficult to obtain, if not at all impossible, in the process for locating the best KPIs. This could be even more problematic when the dynamics of how performance is driven is understood differently within the same organization. For instance, when products are delivered to end users promptly (which is often regarded by sales and marketing as a measure of order completion performance), they are less likely to be accurately meeting the needs in terms of quality (and thus resulting in quality control becomes more challenging for the production team), and the more likely the operation will become costly (Caplice & Sheffi, 1994). To this end, the production manager must decide in advance how he or she wants the KPIs to align along and balance among the three competing attributes of speed, accuracy, and cost, and understand how these selected KPIs could possibly be affected by other factors not directly under his or her control, product attributes, for example. In fact, while the quest for perfection in any manufacturing business has never slowed down since the term Zero Defects was first coined by Philip Crosby in the 1960s upon which the modern day Six Sigma concepts are based on (Suarez, Gera, 1992), contemporary studies and theories, especially the very popular quality management standards proposed by the International Organization for Standardization (ISO) are more geared toward concentrating on the seven quality management principles[2] and have rarely touched on the practical selection of KPIs, let alone the control of product attributes because these quality management principles were not originally designed and tailored for small manufacturing businesses.

[1] Note that there is also previous literature that suggests the other way round in which new KPIs are not necessarily threatening, see Neely et al., (1997; 2000), Stiles et al. (1997), and Viken, (1995).

[2] The seven quality management principles are: customer focus, leadership, engagement of people, process approach, improvement, evidence-based decision making, and relationship management (International Organization for Standardization, 2015)

Revisiting the Cascade

As business objectives and measures that are derived from a firm's core strategy should normally be cascaded through the various levels of an organization to help maximize and achieve performance (Assiri, Zairi, & Eid, 2006, p.946; Kaplan, 2005, p.41; Moullin et al., 2007), it is not unusual for production managers, especially those who are working for a small manufacturing business, to be seen as among those who are stranded in the 'downstream' of such a cascade when business development, sales and marketing, and finance are all seemingly to be in the upstream[3]. Therefore, in order to understand how factors from the upstream (in terms of product attributes) are shaping KPIs in the downstream, an investigation of a list of standardized, easy-to-measure KPIs was carried out at a relatively small-scale, traditional 500-worker metal stamping factory[4] in China over a period of 9 months, from June 2014 to February 2015[5]. A metal stamping factory in China was selected for this research because while there are approximately 20,000 factories in China and roughly some 1,500 are metal stamping factories, most of them, if not all, are holding at least one kind of quality management certification, e.g. ISO 9001, ISO 9000, ISO14001, ISO 14000, ISO 20000 (Made-in-China.com, 2015) yet many of them are still facing issues on quality and defects. These metal stamping factories, unlike their footwear or apparel counterparts in China, are mostly SMEs with less than 500 workers (see Table 1 below) and are, hence, suitable for this research as our focus is on small factories that are not only facing the 'substantial barriers to strategic performance management' (Hudson, Smart, & Bourne, 2001) but also those in which 'actionable knowledge' could hardly be acquired (Meyer, 2005), even when they are mostly ISO-certified.

Table 1. Metal factory size in the China (Made-in-China.com, 2015)

Factory size	Distribution (%), N ≈ 1500
Less than 5 employees	5%
5 to 50	45%
51 to 200	40%
201-500	8%
More than 500	2%

Methodology

The investigation involved 98 shipments of approximately 250,000 pieces of finished metal work products. The defect rate, productivity, OEE (overall equipment efficiency), worker efficiency, and yield loss in these product batches were computed and recorded. These were then compared, using the t-test, with various generic product attributes such as product size, complexity (and hence allocation of manpower and production time), pricing and order types (e.g. new vs. repeat orders, urgency of order, etc.) and these product attributes were then further grouped into several clusters showing similar characteristics according to their respective defect rates and yield losses using k-means clustering and then, finally, revalidated using ANOVA.

Findings

Presuming that the production process has been duly controlled (as this is one of our major research limitations in a non-ideal production setting), the results have been rather fascinating. Regardless of the product model (old or new) and product size, significant differences were found across the respective defect rates ($t=-3.85$, $p<.00$)(see Table 2 below), yield losses ($t=-1.46$, $p<.01$), and even finished goods inventory ($t=-3.27$, $p<.00$) as purchasing period varies (or order urgency changes). Logically speaking, what is interesting is that under normal rationale or thinking, with purchasing periods ranging from two weeks (14 days) to more than two months (60 days), production managers would normally expect that the shorter the purchasing period, the higher the defect rate. Yet in our real-life case study, defect rates actually seem to have been 'contained more effectively' with shorter purchasing periods (even though they do spike a bit as purchasing periods become too short and the orders might be too urgent to handle). And they fluctuate more (and become more difficult to contain) as purchasing periods increases (or urgency of orders decreases), see Figure 1a. This seems to imply that a reasonably short (but not too short) purchasing period (e.g. two weeks) could be the 'optimal purchasing period' for controlling and containing defect rates for the production line – a scenario that might suggest that small-scale factories like the one we have investigated tend to work more efficiently with shorter cash-to-

[3] Small businesses are generally operating in the 'survival mode' and hence the upstream controls everything, including the pricing, the development of the product, etc.

[4] Established in 1993, Best Ideal Limited is an ISO9001:2008, ISO14001:2004, and ISO 13485:2003 certified metal works factory with a factory floor area of over 400,000 square feet in Dongguan, China. The company received it 'Partner in Excellence Award' from Celestica in 1999, 'Partner's in Performance Award' from Celestica for three consecutive years (2002 to 2004), 'Most competitive Supplier Award' from Polycom in 2000, 'Outstanding Supplier Award' from CoVi Technologies in 2005 and 2006, 'TCOO™ Supplier Awards – Most Technical Compliant Award' from Celestica in 2007 and 2008, ADVA Supplier Award 2013 for Innovation, and Best Partner Award from Citi Commercial Bank in 2012. For details, see http://www.bestideal.com/

[5] In order to avoid sensitive issues regarding the data, the data was only released in February 2016, a year after the actual collection.

cash cycles (e.g. approximately two weeks in length). It may also suggest that as the purchasing period gets longer, defects simply become more difficult to control and hence cannot be reasonable contained as a result of various other factors, human related maybe, motivation issues, worker morale etc. However, it has to be emphasized that the traditional learning curve still applies as we can see when order quantity increases, defect rates generally shrinks (see Figure 1b). This is even more obvious when producing the old models, see Figure 2a and Figure 2b. The actual products are also shown in Figure 1c below.

Table 2. Differences in defect rates with different purchasing periods (days)

Days	***T***	**df**	**Sig.**	**Days**	***t***	**df**	**Sig.**
8	0.03	96.00	0.98	44†	4.57	96.00	0.00
	0.05	7.94	0.96		2.72	16.06	0.02
12†	1.99	96.00	0.05	48†	4.57	96.00	0.00
	3.26	85.43	0.00		2.72	16.06	0.02
16‡	3.31	96.00	0.00	52‡	5.00	96.00	0.00
	3.62	74.08	0.00		2.82	13.73	0.01
20‡	3.93	96.00	0.00	56†	4.11	96.00	0.00
	3.54	50.76	0.00		2.41	10.60	0.04
24‡	4.41	96.00	0.00	60	3.67	96.00	0.00
	3.52	39.05	0.00		1.94	8.31	0.09
28‡	4.19	96.00	0.00	64	3.67	96.00	0.00
	3.33	35.03	0.00		1.60	5.09	0.17
32‡	4.33	96.00	0.00	68	1.58	96.00	0.12
	3.06	25.24	0.01		1.28	3.16	0.29
36†	3.99	96.00	0.00	72	1.58	96.00	0.12
	2.63	21.05	0.02		1.28	3.16	0.29
40‡	4.58	96.00	0.00				
	2.89	18.54	0.01				

Figure 1a. Defect rates with different purchasing periods

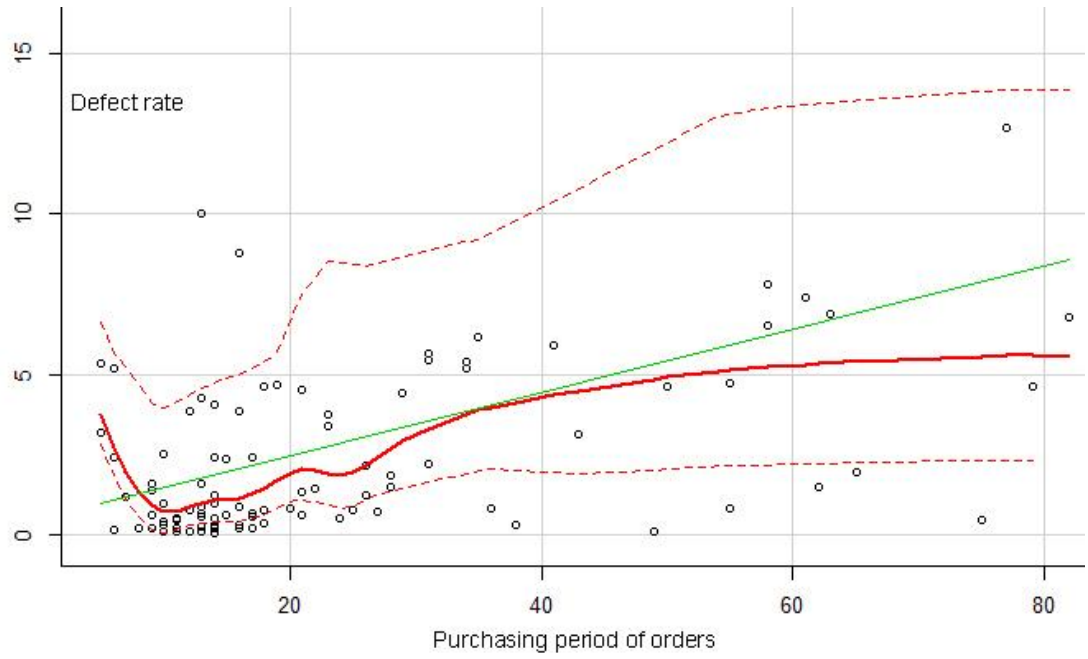


Figure 1b. Defect rates with order quantity for each shipment (pieces)

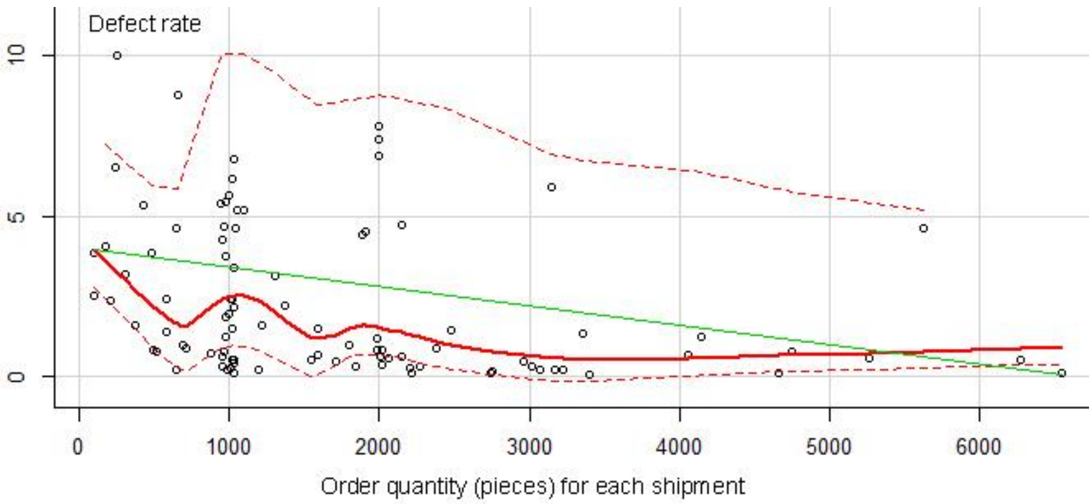


Figure 1c. Photos of actual products[6]

[6] Legend: LPS/MPS = less/more production steps, LS/SS = large/small size, NM/OM=new/old model

Figure 2a. Defect rates with purchasing periods for old and new models

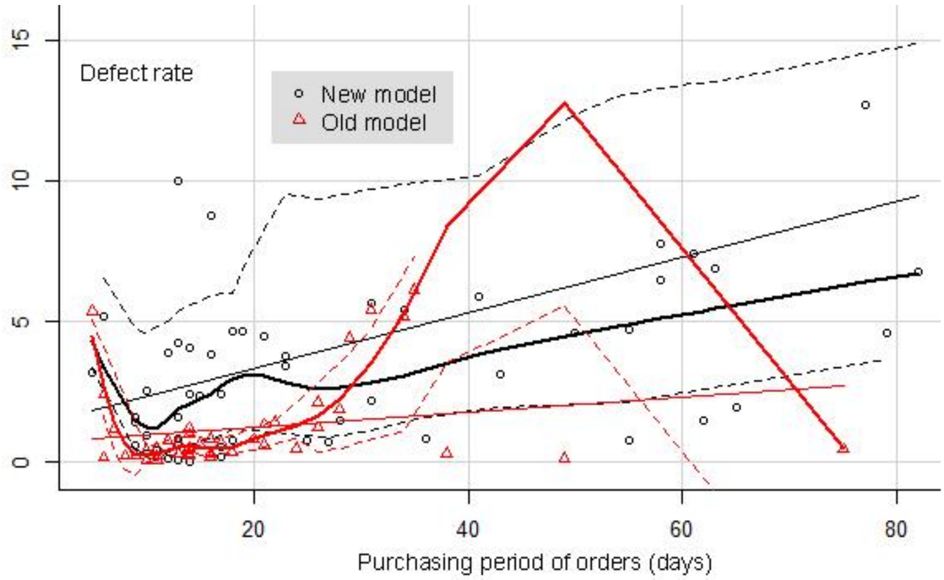
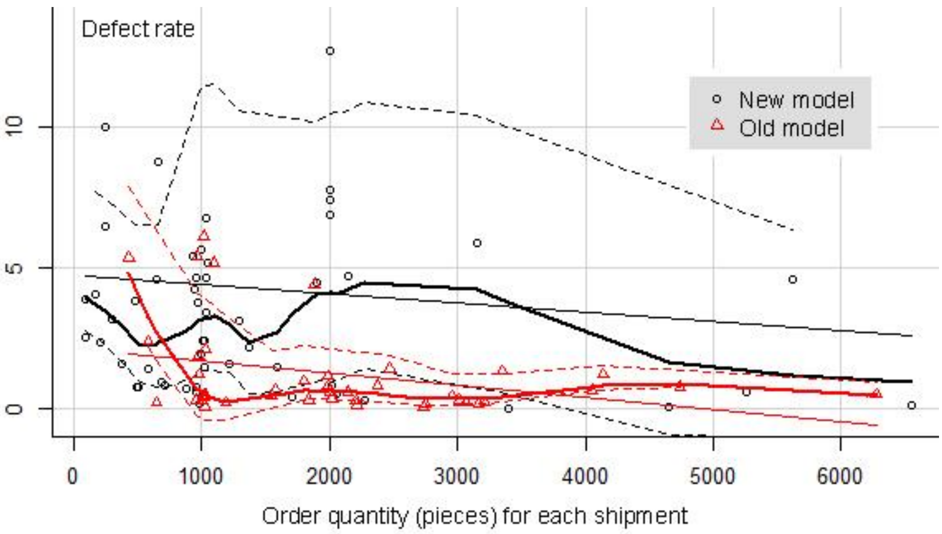


Figure 2b. Defect rates with order quantity for each shipment (pieces) for old and new models



With products involving different levels of difficulty (or complexities), i.e. less versus more production steps (in which 3 or more production steps is considered to be 'more' and hence complex products), defect rates ($t=-2.41$, $p<.00$) and yield losses ($t=-2.40$, $p<.02$) both differ significantly with purchasing period (or order urgency). And between large and small product sizes, significant differences appeared in finished goods inventory ($t=5.36$, $p<.00$), incoming material inventory ($t=-2.63$, $p<.00$), and, most surprisingly, in the overall equipment efficiency (OEE) ($t=-1.05$, $p<.04$), a KPI which is thought to be relatively inert in the production line. And in all three cases, productivity and worker efficiency remains basically unaffected as expected and no significant statistical difference was found among all the different batches.

Figure 3a. Defect rates with purchasing periods for simple and complex models

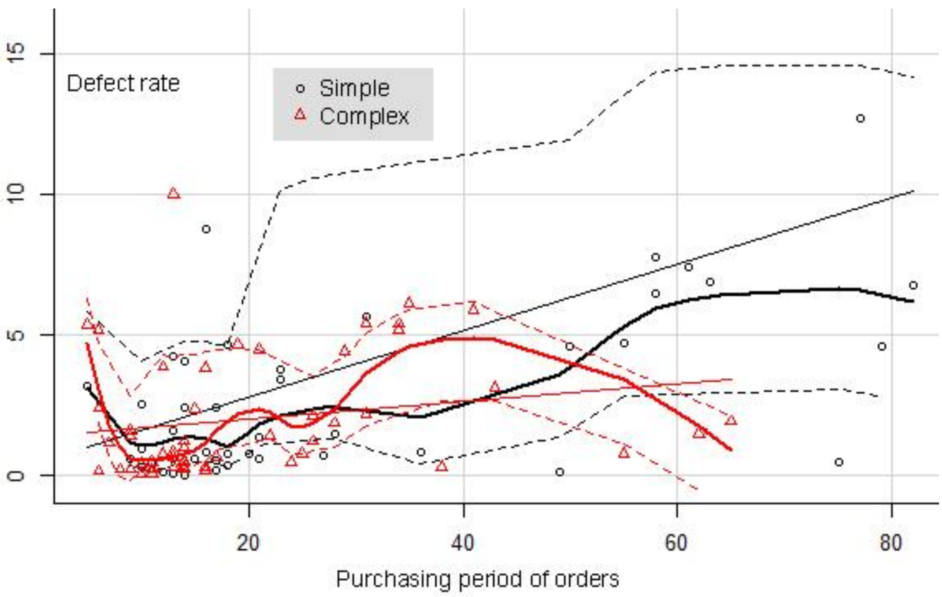
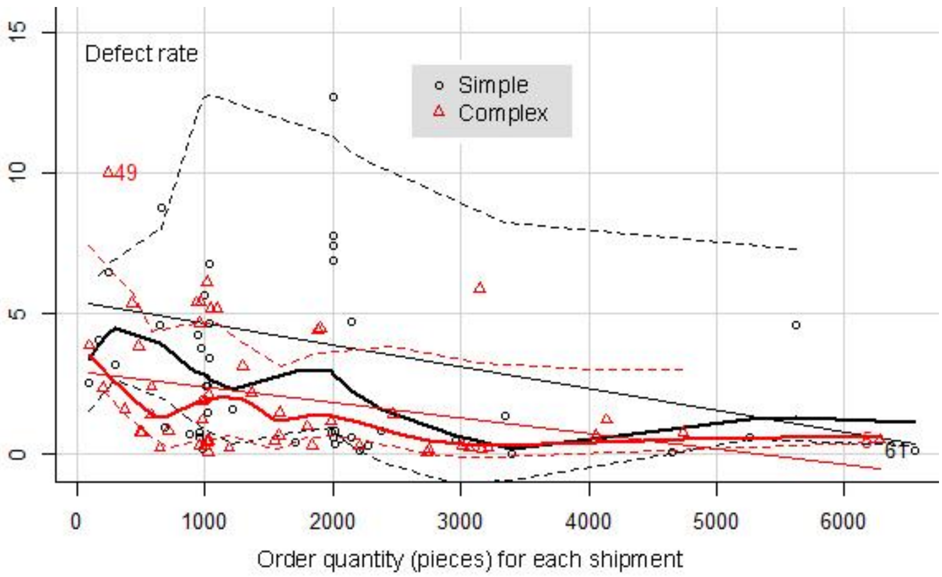


Figure 3b. Defect rates with order quantity for each shipment (pieces) for simple and complex models



Yet with a closer look at these figures (see Figures 3a, 3b, 4a and 4b), one would notice that even though defect rates all shrinks generally as quantity increases in each batch regardless of product complexity and sizes, the longer the purchasing period for a simple product, the harder it is to control the defect rates. Whereas for complex products (i.e. products with more production steps), it seems to be relatively more controllable, if not entirely opposite – the longer the purchasing period, the more controllable it may seem with the defect rates. But since there are practically few orders (about 16 out of 98 batches) that would allow for such unnecessarily lengthy purchasing time, our findings for the defect rates with exceptionally long purchasing periods (e.g. over 2 months) may seem to be inconclusive.

Figure 4a. Defect rates with purchasing periods for different product sizes

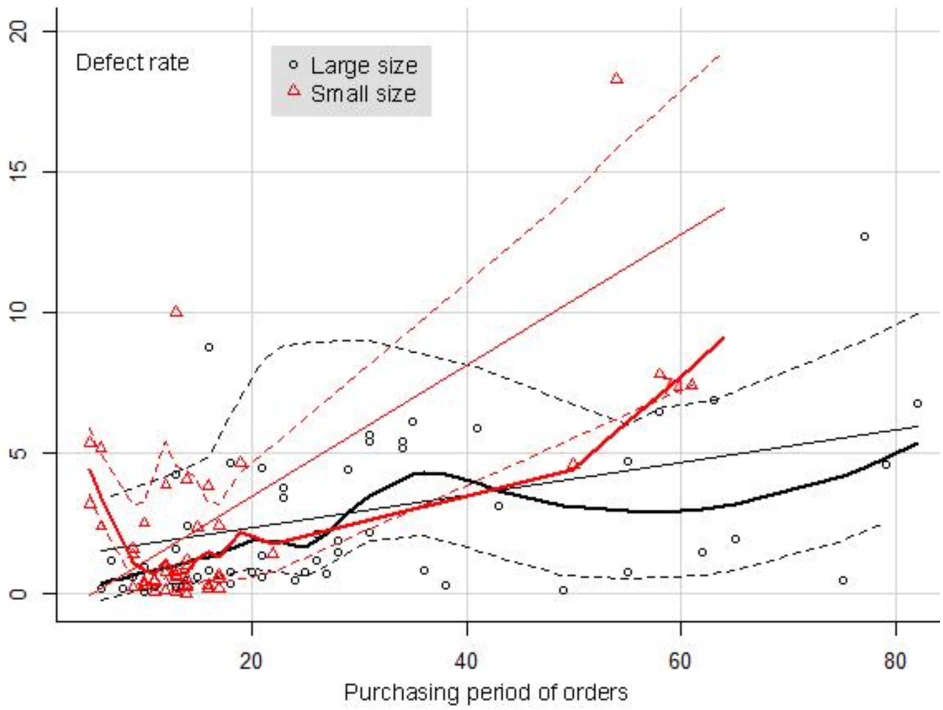
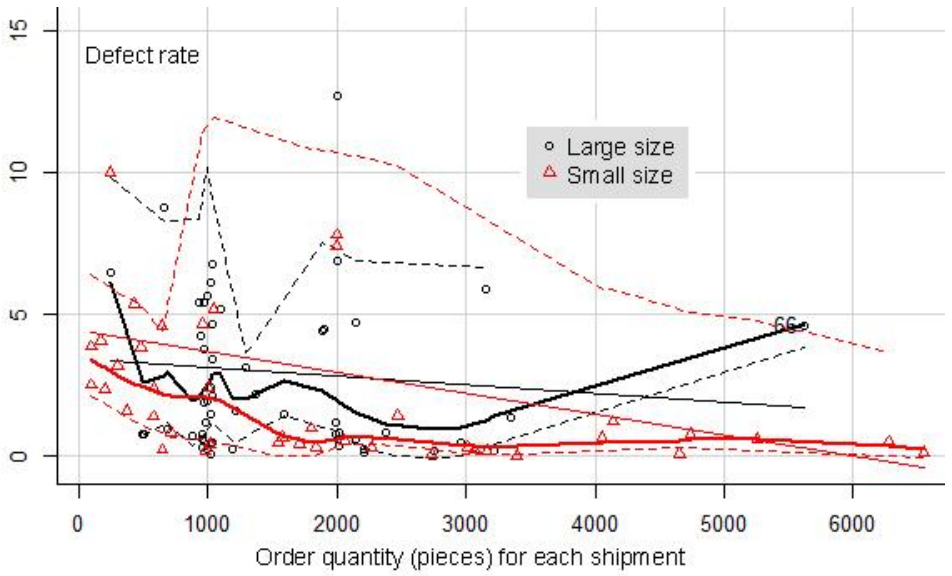


Figure 4b. Defect rates with order quantity for each shipment (pieces) for different product sizes



Further Exploratory Findings
Theoretically speaking, under any given production setting, any combination of product sizes, model, complexity, quantity and purchasing periods should be possible. Yet when it comes to practical consideration, only a number of them actually exist. So, in order to study what kind of combination exist in a practical production setting the 98 batches of products were further divided into 4 clusters (using k-means clustering) according to their respective attributes including product size, quantity, complexity, order types (e.g. new vs. repeat orders, urgency of order) and defect rates. Some very interesting patterns showed up, see Table 3a below.

Table 3a. Production batches clustered (with mean values) according to attributes and defect levels[7]

Cluster	N	Size	Model	Complex	Quantity	Urgency	Defect
1	14	1.07	1.86	1.71	1.07	1.21	2.93
2	25	1.80	1.84	1.32	1.76	2.76	3.64
3	29	1.90	1.55	1.34	1.69	2.21	1.38
4	30	1.27	1.23	1.80	2.33	1.13	1.10
Total	98						

Table 3b. ANOVA table revalidating the 4 product clusters

[7] Product sizes are categorized as either 1 (small) or 2 (large); model types as either 1 (old) or 2 (new); complexity as 1 (simple, less than 3 production steps) or 2 (complex, more than 3 steps); quantity values range from 1 (under 1000 pieces), 2 (from 1000 to 3000 pieces) to 3 (over 3000 pieces); order urgency values range from 1 (within 2 weeks), 2 (2 to 4 weeks) to 3 (over 3 weeks); defect levels, instead of rates, range from 1 (below 1%), 2 (from 1% to 2.5%), 3 (from 2.5% to 5% of defects) to 4 (over 5%).

		Sum of Squares	df	Mean Square	F	Significance
factor_size	Between Groups	10.65	3	3.55	24.74	.00
	Within Groups	13.48	94	.14		
	Total	24.13	97			
factor_model	Between Groups	6.39	3	2.13	11.36	.00
	Within Groups	17.61	94	.19		
	Total	24.00	97			
factor_steps	Between Groups	4.76	3	1.59	7.59	.00
	Within Groups	19.65	94	.21		
	Total	24.41	97			
factor_quantity	Between Groups	16.33	3	5.44	18.04	.00
	Within Groups	28.36	94	.30		
	Total	44.69	97			
factor_purchase	Between Groups	45.39	3	15.13	61.45	.00
	Within Groups	23.14	94	.25		
	Total	68.53	97			
factor_defect	Between Groups	113.96	3	37.99	160.72	.00
	Within Groups	22.22	94	.24		
	Total	136.17	97			

From Table 3a above, we can see when the product size is relatively small (1.07 and 1.27) and can easily be handled (as in clusters #1 and #4), these are always complex orders that require more production steps (1.71 and 1.80) and the defect levels go up (from 1.10 to 2.93) with newer models (from 1.23 to 1.86). And, maybe, because of various market risks in introducing new model products, order quantities for newer models are always smaller; hence further pushing up the defect rates. Older models or repeat orders (as in cluster #4), on the contrary, come with much larger average order quantities (2.33) and hence allow defect rates to go down as a result of the learning curve, even when purchasing periods are short (1.13). However, when the product size gets bigger (as in cluster #2 and cluster #3), production schedules from the real-life cases become much more flexible and more average time is usually allowed (2.21 and 2.76) for the manufacturing of the batches even when their complexity is relatively low on average (1.32 and 1.34). Yet the defect rates, for the cases of these large products, seem to go up again with the model. Newer models experience much higher defect rates (3.64) whereas the models score much lower, better rates (1.38).

Conclusion

Practical production, especially in China where personnel issues are still prevalent, is sometimes different from theoretical calculations, especially when most of the leading theories in the field are focused upon the large enterprises with thousands of employees, if not more. In this paper, we set off to search for better combinations of Key Performance Indicators (KPIs) by studying a list of easy-to-measure KPIs compared with the attributes of the respective products. Our results show that while effective quality production management should theoretically be able to help minimize or control unwanted variations in defects and yield losses, product attributes such as size, complexity and hence the urgency of orders do come in some particular combinations (as shown in our cluster analysis) and these combinations can mean effects over the entire production line that are not immediately visible to the dedicated production manager. If the sales and marketing arm of a manufacturing business could be made aware of such a relationship between product attributes and the corresponding outcome (in terms of defect rates, etc.), these small businesses could easily realize much more effective pricing strategies which could jack up profit margins when now they know that defect rates for a particular product would go up. Conversely, if sales and marketing data could be taken into consideration as new KPIs for the production manager, much better control of the quality of products could as well be anticipated, regardless of the country of production.

References

Assiri, A., Zairi, M., & Eid, R. (2006). How to profit from the balanced scorecard: An implementation roadmap. Industrial Management and Data Systems, 106(7), 937-952.

Caplice, C. & Sheffi, Y. (1994). A review and evaluation of logistics metrics. The International Journal of Logistics Management, 5(2), 11-28.

Decoene, D. & Bruggeman, W. (2006). Strategic alignment and middle-level managers' motivation in a balanced scorecard setting. International Journal of Operations and Production Management, 26(4), 429-448.

Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives [Abstract]. Management Review (AMA Forum), 70(11), 35-36.

Hudson, M., Smart, A., & Bourne, M. (2001). Theory and practice in SME performance measurement systems. International Journal of Operations and Production Management, 21(8), 1096-1115.

Interational Organization for Standardization (ISO). (2015). Quality management principles [PDF version]. [Brochure]. Retrieved from http://www.iso.org/iso/pub100080.pdf

Kaplan, R.S. (2005). How the balanced scorecard complements the McKinsey 7-S model. Strategy and Leadership, 33(3), 41-46.

Keegan, D.P., Eiler, G. & Jones, C.R. (1989). Are your performance measures obsolete? Management Accounting, June 1989, 45-50.
Lambert, D. (2001). Supply chain metrics. The International Journal of Logistics Management, 12(1), 1-20.
Made-in-China.com. (2015). Number of metal stamping industries in China. Retrieved June 1, 2016 from http://www.made-in-china.com/companysearch.do?subaction=hunt&mode=and&code=0&style=b&isOpenCorrection=1&word=Metal+Stamping+Industries&comProvice=nolimit
Meyer, M. (2005). Can performance studies create actionable knowledge if we can't measure the performance of the firm? Journal of Management Inquiry, 14(3), 287-291.
Moullin, M., Soady, J., Skinner, J., Price, C., Cullen, J., & Gilligan, C. (2007). Using the public sector scorecard in public health. International Journal of Health Care Quality Assurance, 20(4), 281-289.
Neely, A., Huw, R., Mills, J., Platts, K., & Bourne, M. (1997). Designing performance measures: A structured approach. MCB International Journal of Operations and Production Management, 17(11), 1131-1152.
Neely, A., Mills, J., Platts, K., Richards, H., Gregory, M., Bourne, M. & Kennerley, M. (2000). Performance measurement system design: Developing and testing a process-based approach. International Journal of Operations and Production Management, 20(10), 1119-1133.
Spitzer, D. (2007). Transforming performance measurement: Rethinking the way we measure and drive organizational success. AMACOM, Division of American Management Association ISBN 9780814408919.
Stiles, P., Gratton, L., Truss, C., Hope-Hailry, V. & McGovern, P. (1997). Performance management and the psychological contract. Human Resource Management Journal, 7(1), 57-66.
Stivers, B. P., & Joyce, R. (2000). Building a balanced performance management system. SAM Advanced Management Journal, Spring 2000, 22-29.
Suarez, J. Gerald. (1992). Three experts on quality management: Philip B. Crosby, W. Edwards Deming, Joseph M. Juran [PDF version]. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a256399.pdf
Viken, K. (1995). Team measurement: Some whys, whats and hows. Retrieved from CSWT Reports, website: http://www.workteams.unt.edu/reports/Viken.html

Raymond Cheng
Reader, Industrial Doctorate (IndD) Programme,
Asia e University, Kuala Lumpur, Malaysia.
Email: raymond.cheng@kellogg.oxon.org

Pit-yan Li
Senior Quality Manager, Best Ideal Ltd, Hong Kong.
Email: hpyli@yahoo.com.hk