



**HAL**  
open science

## Strategic Workforce Planning and sales force : a demographic approach to productivity

Marie Doumic, Mathieu Mezache, Benoît Perthame, Edouard Ribes, Delphine Salort

### ► To cite this version:

Marie Doumic, Mathieu Mezache, Benoît Perthame, Edouard Ribes, Delphine Salort. Strategic Workforce Planning and sales force : a demographic approach to productivity. 2017. hal-01449812

**HAL Id: hal-01449812**

**<https://hal.science/hal-01449812v1>**

Preprint submitted on 3 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Strategic Workforce Planning and sales force : a demographic approach to productivity

Marie Doumic \*      Peter Ingraham<sup>†¶</sup>      Mathieu Mezache<sup>‡</sup>      Benoît Perthame<sup>§</sup>      Edouard Ribes<sup>¶||</sup>  
Delphine Salort<sup>||</sup>

February 3, 2017

## Abstract

Sales force Return on Investment (ROI) valuation with marketing mix frameworks is nowadays common. Sizing discussions then generally follow based on market data and business assumptions. Yet, according to our knowledge, little has been done to embed sales force demographic data (age/experience, tenure, gender etc...) as well as dynamics (especially turnover) in order to investigate the impact of the salesforce characteristics on sales. This paper illustrates such an attempt. It shows that sales force ROI valuation can benefit from a correction on turnover and that optimizing a sales rep hiring policy can unleash additional ROI points. The results are yet heavily dependent in the data structure of the study and their generalization would have to be investigated.

---

2010 *Mathematics Subject Classification.* 90B70, 91D35, 92D25, 35B40

*Keywords and phrases.* Marketing Mix; Learning Curve; Optimal Policies; Sales Force Turnover; Workforce planning; Structured population dynamics

## Introduction

**Goals and Motivations.** On-the-job productivity improvement is a well-known fact conceptualized under the name of learning curve. This first appeared in the seminal work of Wright in 1936 [Wright, 1936]. If the original focus was more of an industrial nature, the concept has nowadays been mirrored and applied to multiple labor areas.

But the productivity increase associated to time and task repetition is challenged by individual labor dynamics, especially turnover. This raises the question of the optimal demographic company composition, as productivity level and worker lifecycles have been empirically proven to be sensitive to such parameters [Ng and Feldman, 2009]. This topic yet remains a broad challenge because of the

---

\*Inria de Paris, EPC Mamba, UPMC et CNRS, F75005 Paris, France

<sup>†</sup>Email: peter.ingraham@yahoo.com

<sup>‡</sup>Sorbonne Universités, UPMC Univ Paris 06, Laboratoire Jacques-Louis Lions UMR CNRS 7598, Inria, F75005 Paris, France

<sup>§</sup>Sorbonne Universités, UPMC Univ Paris 06, Laboratoire Jacques-Louis Lions UMR CNRS 7598, Inria, F75005 Paris, France

<sup>¶</sup>Email: edouard.augustin.ribes@gmail.com

<sup>||</sup>Sorbonne Universités, UPMC, Laboratoire de Biologie Computationnelle et Quantitative UMR CNRS 7238, F75005 Paris, France

difference in productivity definition and work cycles across labor categories (notably through the central distinction between direct and indirect workforce, direct workforce being directly linked to the production, contrarily to indirect one) as well as because of its sensitivity to work organizations, *e.g.* company with high tenure culture.

In [Doumic et al., 2016], we considered that a company labor supply is primarily driven by turnover. An age-structured population model was therefore built. It revealed efficient to answer workforce planning issues, such as how to hire new workers optimally to both mitigate the demand-supply gap and minimise the costs. This new study goes a step further: instead of focusing on the management of the labor cost, a productivity optimisation is proposed. On top of that, the set of demographic variables used to describe the workforce has been extended. Workers age, tenure and gender have indeed been employed to better capture possible learning curve effects on productivity.

A case study on a pharmaceutical salesforce in an emerging market is hereby proposed. Sizing and staffing such a workforce is traditionally under special scrutiny. Sizing can be well managed through promotion response analysis. This direct population indeed constitutes an easy application target for this kind of analysis because of the abundance of data associated to Customer Relationship Management (CRM) systems.

This study starts by measuring the influence of workforce activity on pharmaceutical sales by the means of a dynamical model (Section 1). It provides a productivity argument to the sizing of such a workforce. Then the effects of demographic characteristics (such as gender, age and tenure) on turnover and productivity are investigated in Section 2. In Section 3 finally, we embed the demographic information found in the previous sections into a dynamic sales model to analyse the productivity sensitivity to both the sizing and the demographic composition of the workforce. The model is an age and tenure-structured equation, given by the system (17)–(19). To our knowledge, little has been done to standardize the quantitative impact of staff composition under an overall productivity standard.

**Important Legal Remarks.** The findings and opinions expressed in this paper are those of the authors and do not reflect any positions from any company or institutions. Note that for the sake of confidentiality, the market is not specified in this paper. Note that consideration of age, gender and tenure may not be appropriate for all labor markets. For instance, the US and the UK prohibits the differentiation by age in the recruitment process. Applying this paper on such market would require a change in variables. Yet we would expect the results to yield similar results such as the ones developed in the last part of this paper, which revolve around a comprehensive quantitative framework that provides a ROI - optimal argument for pharmaceutical sales representatives hiring profile. Finally, please bear in mind that to protect confidentiality numbers have been disguised in a way that preserves the same analysis and conclusions as the actual case study.

## 1 Sales Force ROI through Promotion Response Analysis?

Product diffusion and response to promotion is a well-known topic in marketing research (see [Peres et al., 2010] for a recent review). Yet published applications in the pharmaceutical industry are rather sparse [Stremersch, 2008]. In this domain, promotional tools mainly fall under two categories: direct to physician and direct to consumer advertising ([Calfee et al., 2002]). Their respective effectiveness can drastically change depending in their field (therapeutic area) and market of application (see

[Kremer et al., 2008] for a review on pharmaceutical promotion elasticity and [Desiraju et al., 2004] for a discussion on the emerging vs mature market paradigm). Moreover academics do not reach a consensus on this topic in terms of modelling, due notably to the complex nature of the industry - physician - patient interaction. For instance, some of the developed frameworks are extremely close to the standard Bass innovation diffusion model [Vakratsas and Kolarici, 2008], while others match the notion of repeated purchase associated to chronic conditions ([Hahn et al., 1994]) or encompass marketing mix instruments considerations ([Lilien et al., 1981]).

### 1.1 Sales Data Description.

The product described in this study belongs to a consumer healthcare category. The country associated to the analysis can be considered as an emerging market. Note that to protect confidentiality numbers have been disguised in a way that preserves the same analysis and conclusions as the actual case study. Sales data (in product standard units) are described over a period of 24 months over a total of about 3700 bricks (geographical sub units). The associated report also enabled differentiation of the sales units among 3 main packaging forms (referred below as *A*, *B* and *C*).

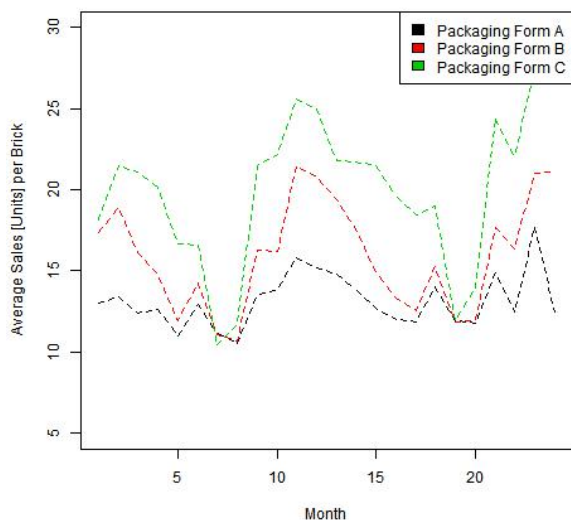


Figure 1: Seasonal Patterns in Sales evolution

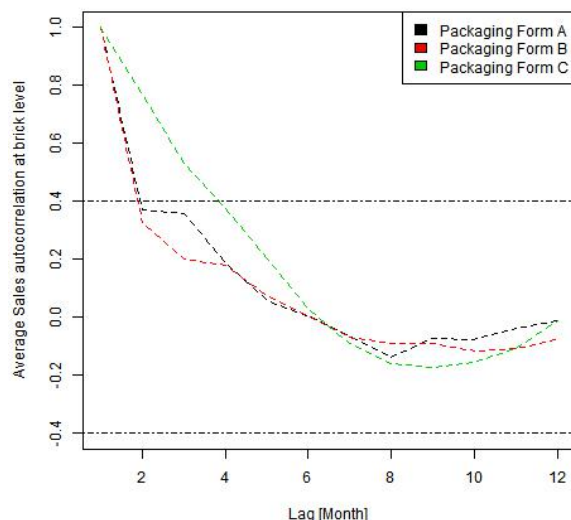


Figure 2: Autocorrelation in Sales evolution

The 3 main packaging forms exhibit different patterns. The *B* and *C* forms show a strong seasonality pattern over the same period of time which is not the case for the form *A*. Form *C* also exhibits an autocorrelation up to 4 months, while forms *A* and *B* show an autocorrelation up to 2 months.

The detailing activity of the pharmaceutical salesforce associated to the product has been reported at brick level with monthly granularity for both the same geographical and time span. The detailing activity report also explains which physician (along with his speciality) each of the salespersons has reached and how many calls have been made. This type of granularity allows to estimate the detailing sensitivity to physician type for optimization purposes. The salesforce in scope of this study consisted of about 250 salespersons in a "polyrep" model (1 salesperson promotes multiple products). The physician specialities were described at a macro and micro level (ex: surgeon (macro level), pediatric

surgeon (micro level)). About 60 micro speciality categories were available. Note that the product prices at pharmacy level for the 3 packagings forms were available at brick level as well as the average price of the competing products.

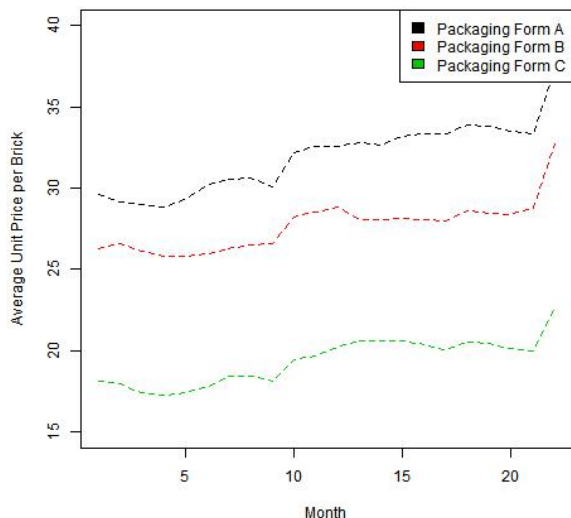


Figure 3: Pharmacy level Price Evolution

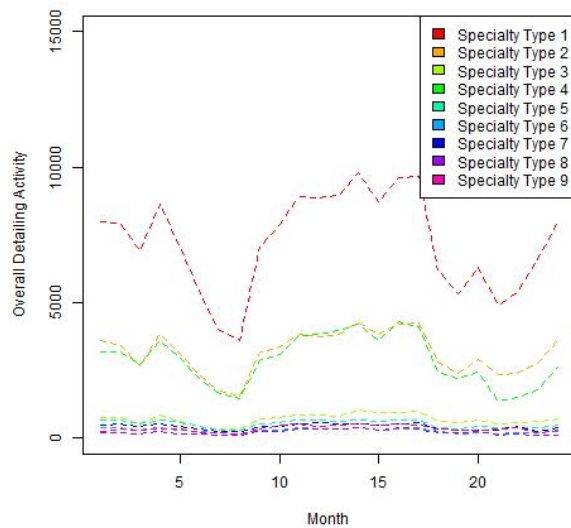


Figure 4: Seasonal Patterns in Detailing Activity (number of calls)

3 out of these micro specialities concentrated 78% of the overall detailing effort, an additional 6 accounted for a remaining 15% of the overall detailing activity, while the rest of the micro speciality could be neglected. Also note that the detailing activity is extremely seasonal.

## 1.2 Marketing Mix Framework and Calibration Results.

**Marketing Mix Framework.** Pharmaceutical product sales are usually thought to consist of carry-over sales (*i.e.*, the sales linked to the *past* selling activity which would occur with no active *present* selling activity) and impactable sales (*i.e.*, the sales directly reflecting promotional activities) [Cook, 2015]. To our knowledge, carry-over rates stem either from customer loyalty or prescription refill, while impactable sales translates the direct effect of promotional tools such as detailing, journal advertising, etc.

The prominence of carry-over, especially for drugs related to chronic conditions as the one depicted in this analysis, suggests a time-series analysis approach. This raises the question of the lags to employ in the model [Manchanda and Chintagunta, 2004]. Standard cross-correlation tests were therefore used on the available data to select an order for the autoregression part of the model. The order was then adjusted to achieve a best-fit according to the Akaike information criterion (AIC).

For the impactable sales part of the model, three main factors where taken into account: seasonality, detailing activity and price changes. Usual parametrisations for the dependence of the sales in these factors vary from exponential to linear and log forms. In our case study, log forms similar to [Leonard Jon Parsons, 1981] were used. A carry-over structure was used to describe sales evolution to account for the observed autocorrelation in sales. This translates into the following simple framework,

where  $i$  designates the brick level analysis,  $f = A, B, C$  the product formulation,  $t$  the time of the observation,  $p_{i,t,f}^{Brand}$  the product price at time  $t$  for form  $f$  and on brick  $i$  and  $p_{i,t,f}^{Competition}$  the average price of the competition for the product at time  $t$  for form  $f$  and on brick  $i$ :

$$\log(\text{sales}_{i,t,f} + 1) = \underbrace{\sum_{q=1}^{q=4} Q_{q,f} \cdot 1_{t \text{ in quarter } q}}_{\text{Quarterly Seasonal Sales level}} + \underbrace{\sum_{o=1}^{o=O} \log(\text{sales}_{i,t-o,f} + 1) \theta_{o,f}}_{\text{carryover sales}} + \underbrace{\beta_f \log(D_{i,t} + 1)}_{\text{detailing impacted sales}} + \underbrace{\lambda_f^{brand} \log(p_{i,t,f}^{brand} + 1)}_{\text{Brand price impacted sales}} + \underbrace{\lambda_f^{Competition} \log(p_{i,t,f}^{Competition} + 1)}_{\text{Competition price impacted sales}} + \underbrace{\epsilon_{i,t,f}}_{\text{Model error}} \quad (1)$$

In the equation 1,  $Q_{q,f}$  reflects the basal sales level in quarter  $q$  for form  $f$ ,  $\beta_f$  represents the sales sensitivity to the detailing activity  $D_{i,t}$  and  $\lambda_f^{brand}$  (resp.  $\lambda_f^{competition}$ ) is the sensitivity of sales to the product (resp. competition) price for form  $f$ . Note that due to the granularity of the available data, a hierarchical bayesian framework could be employed to further explore the errors terms and dependencies of the various contributions to the sales.

**Calibration To Market Data.** The overall model is performing well on all 3 packaging forms ( $R^2 > 90\%$ ). The performed regression (based on the data previously described) resulted in the following parameters values (standard deviation (std) in parenthesis):

Coefficient	Form A	Form B	Form C
$Q_1$	-	6.52 (0.22)	6.71 (0.11)
$Q_2$	-	6.61 (0.23)	7.038 (0.12)
$Q_3$	-	6.70 (0.22)	7.26 (0.11)
$Q_4$	-	6.42 (0.22)	6.63 (0.11)
$\theta_1$	0.65 (0.002)	0.66 (0.002)	0.46 (0.003)
$\theta_2$	-	-	0.09 (0.004)
$\theta_3$	-	-	0.17 (0.003)
$\beta_1$	0.10 (0.003)	0.10 (0.003)	0.07 (0.002)
$\lambda_f^{brand}$	-0.063 (0.017)	-1.17 (0.07)	-2.46 (0.04)
$\lambda_f^{competition}$	0.44 (0.02)	-0.40 (0.01)	0.65 (0.02)
$R^2$	0.93	0.94	0.95
$\sigma_{t,f}$	0.71	0.74	0.70

As expected the product shows statistically significant seasonal variations for the  $B$  and  $C$  forms. Among the various forms, the sales dependencies toward detailing ( $\beta_1$ ) are low compared to the competition price ones ( $\lambda_f^{competition}$ ). Sales benefits from product price reduction as indicated by the negative signs in  $\lambda_f^{brand}$ . This was expected because of the consumer health care [CHC] nature of the product. CHC products indeed don't have the same access constraints nor the same pricing and reimbursement schemes than traditional prescription medicines. As such patients can in most cases buy products without physician scripts and are provided with multiple options. Since they are often not fully reimbursed, their sensitivity to price is increased. Similar interpretations yield for the sales sensitivity to the competition price with an exception on form  $B$ . Note that this type of features tends to suggest a volume based market strategy.

### 1.3 Sale Force Activity and pricing policies ROI

**Valuation Framework.** Marketing mix models are important, because they allow an optimization of marketing policies and promotional activities. This is usually done by estimating the long term ROI of an investment or a change over a time horizon  $H$  of 3 to 5 years [Andris A. Zoltners, 2006]. In the case described above, two main possible changes can be evaluated: a change in price  $P_f$  of the form  $f$  and a change in the call activity of the sales force  $D$ . The overall sales profit is then given by  $\Pi$ . Note that, for the sake of simplicity, price, competition landscape and detailing activities are assumed constant over the time of the valuation, while estimates are averaged at brick level.

$$sales_{t,f} = \mathbb{E}(e^{\log(sales_{i,t,f}+1)} - 1) = e^{\sum_{q=1}^{q=4} Q_{q,f} + \sum_{o=1}^{o=O} \log(sales_{i,t-o,f}+1)\theta_{i,o,f} + \beta_{i,f} \log(D_{i,t}+1)} \cdot e^{\lambda_{i,f}^{brand} \Delta \log(p_{i,t,f}^{brand}+1) + \lambda_{i,f}^{Competition} \Delta \log(p_{i,t,f}^{Competition}+1)} \cdot \mathbb{E}(e^{\epsilon_{i,t,f}}) - 1 \quad (2)$$

$$\Pi = \sum_f \sum_{t=0}^{t=36} P_f \cdot sales_{t,f}(P_f, D) \quad (3)$$

**Price Change - ROI.** Changing the price  $P_f$  of the form  $f$  by  $x\%$  yield a ROI of:

$$\text{Price Change ROI} = \frac{\Pi(P_f \cdot (1+x), D)}{\Pi(P_f, D)} - 1 \quad (4)$$

As previously described, lowering the prices and promoting a volume based strategy yields a strong ROI under the current assumptions. Note that all forms are not equal. The A form indeed shows little profit variation to a price change because of its inelastic features. But if such models are providing interesting insights on a possible price position, they bear risks. The first one is associated to the market dynamics and the competitors position. The oligopoly (if not monopoly) nature of the pharmaceuticals market would indeed need to be better described. A strong limitation to such studies also lies in the number of prices and intermediary (wholesalers, pharmacists etc...) that exists on the product value chain. The multiple prices indeed affect the overall consumption pattern. Finally the market cap inherent to the prevalence of the disease is not described here. All those features should directly impact the model response structure and would need to be further investigated.

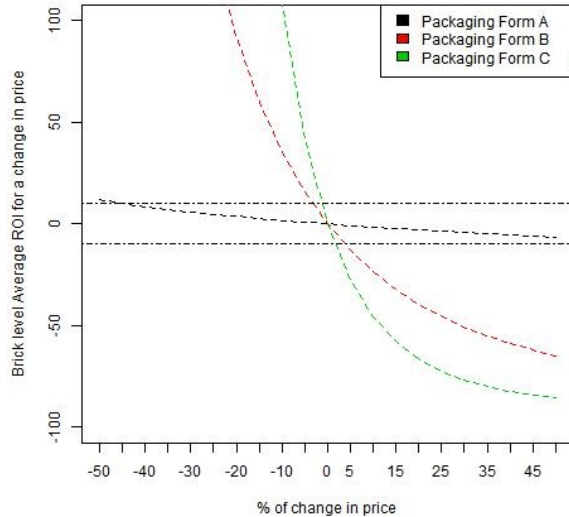


Figure 5: Price Change - 3 years ROI

**Detailing activity shift - ROI.** Changing the call activity  $D$  of the sales force by  $x\%$  and assuming a call unit cost of  $C_d$  yields a ROI over the 3 year period  $T$  of:

$$\text{Call Activity Change ROI} = \frac{\sum_f \Pi(P_f, D \cdot (1+x)) - T \cdot C_d \cdot D \cdot (1+x)}{\sum_f \Pi(P_f, D) - T \cdot C_d \cdot D} - 1 \quad (5)$$

To estimate  $C_d$ , the average salary of a pharmaceutical sales representative over a year was divided by its average number of calls for the period and by its average working days. This resulted in the figure displayed on the right.

Because of the relatively small number of calls at brick level, the graph on the right shows only a small ROI increase for a small increase in call activity. As calls have an integer nature, the continuous framework valorization used here reaches its limits. The general pattern yet states that it is interesting to increase the call activity. Note that the call activity change ROI exhibits the expected decreasing return structure.

Such insights can easily be used to generate a preliminary discussion around the salesforce size and its overall ROI. But they are limited in the way the salesforce cost is described as well as in the normalisation of the call activity. Further investigations in the call response structure and costs should be carried out.

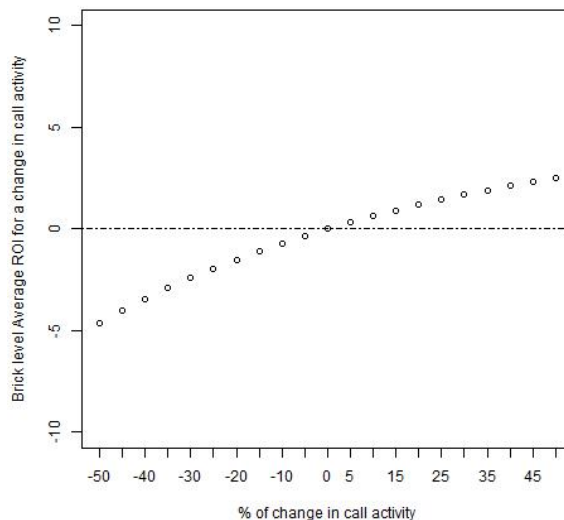


Figure 6: Call Activity Change - 3 years ROI

In this section, a dynamic model of pharmaceutical sales was proposed, according to literature standards, in order to stress the dependency of sales to promotional activity (expressed in terms of call volumes). Estimating its parameters has enabled the computation of an optimal promotional activity level according to ROI criteria. This optimal call volume at country level can then easily be translated into a salesforce size.

The current model assumes that all calls have the same impact, while the ROI valuation assume that the amount of calls at brick level stays constant over time. But two factors can be used to challenge those assumptions. First salesforce turnover pertubates the call volume and turnover is known to be influenced by the salesforce demographics. Second, call impact on sales can change depending on the salesperson that is calling the physicians. We therefore propose to investigate the order of magnitude of those phenomenas by leveraging demographic information on the salesforce. This will help determine wether or not a demographic optimization of a sales force is necessary.

## 2 Is there a need for Sales force demographic Optimization?

Salesforce optimization is a well-known topic. Salesperson hiring profiling has already been investigated in order to achieve optimal long-run performance [Darmon, 1982]. This type of research has yet stressed the need to address the entire salesforce lifecycle within the company as an optimization. This has translated into numerous research regarding salesforce turnover and the associated costs [Darmon, 1990] and has triggered numerous discussion regarding the salesforce segmentation covariates. Demographics characteristics (age [Lucas et al., 1987], tenure, gender[Muchinsky and Morrow, 1980]) as well as socioeconomics characteristics [Karen E. Flaherty, 2002] have commonly been used to perform such optimization ([Darmon, 2004]). Classical frameworks stressed the prominence of direct costs (separation cost, recruiting costs, training costs) over indirect ones (territory vacancy costs, differential operating costs, differential skill costs)[Darmon, 2004]. It translated into interesting salesforce sizing



revisions toward salesforce oversizing or undersizing depending on operational constraints in order to reach maximum sales efficiency.

Such frameworks have mainly focused on direct costs and do not integrated the impact of salesforce socioeconomics characteristics to market dynamics. However in the case of the pharmaceutical industry, this assumption is to be reviewed. Indirect costs especially the cost of a vacant territories may have profound impacts that outwages direct costs. On top of that, the pharmaceutical industry is known to leverage heavily detailing techniques, which translates in important salesforces. Therefore demographic optimization may provide some savings. To our knowledge, those two notions have not been investigated before.

## 2.1 Workforce Data Description.

The previously described commercial data has been completed by an extract from the main Human Resource Information System (HRIS) system associated to the company salesforce. This included salespersons individual demographic information (gender, age and tenure). Age has been reported in bands of 5 years, while tenure has been reported in bands of 3 years. The salesforce showed a relative gender balance during the span of the analysis with about 44% of female. The average salesperson had about 8 years of tenure in the company and was 35 years old. In the company no quantitative guideline was edited regarding the salesperson hiring process and the link between the demographic composition and the sales performance was not investigated. This resulted in arbitrary decisions regarding the salesforce demographic profile to hire.

## 2.2 Salesforce Turnover Impact and Demographic Drivers:

Turnover can be splitted in three categories: voluntary, involuntary and retirement. Voluntary turnover corresponds to employee resigning for personal reasons. Involuntary turnover corresponds to dismissals induced by the company for performance or redundancy reasons.

As such, the turnover rate  $\mu$  can be described over the gender  $g$ , the age  $a$  and the tenure  $t$  as:

$$\mu_g(a, z) = \mu_g^{involuntary}(a, z) + \mu_g^{voluntary}(a, z) \tag{6}$$

Figures 7, 8 and 9 show the turnover rates observed for the salesforce described in this paper.

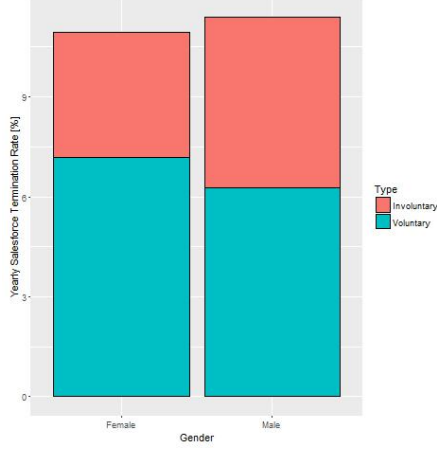


Figure 7: Salesperson Termination Rate with Gender

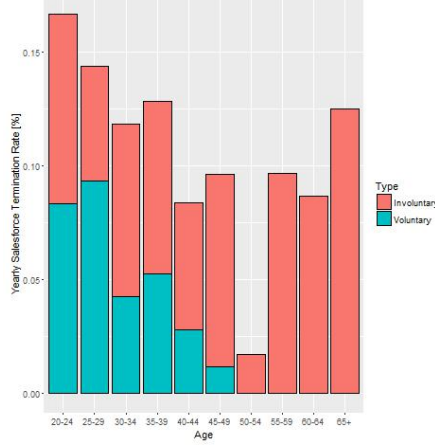


Figure 8: Salesperson Termination Rate with Age

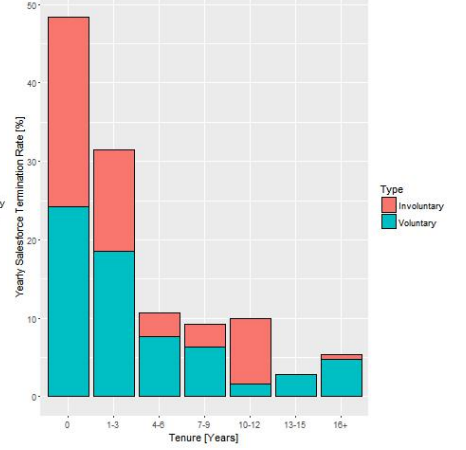


Figure 9: Salesperson Termination Rate with Tenure

At first glance, the overall turnover rate is not really different across genders. (A confusion matrix was indeed created to compare the number of leavers among the salesforce across genders. The  $\chi^2$  test associated this confusion matrix yield the following results:  $\chi^2 = 0.013749$ ,  $df = 1$ ,  $p = 0.9067$ ). Yet Male and Female turnover reasons differ. (A confusion matrix was indeed created to compare the number of voluntary leavers among the salesforce across genders. The  $\chi^2$  test associated this confusion matrix yield the following results:  $\chi^2 = 0.99079$ ,  $df = 1$ ,  $p = 0.3195$ ). Both Voluntary and Involuntary turnover decrease with tenure. Voluntary Turnover decreases with age but involuntary turnover is relatively constant across age bands.

Involuntary turnover can be considered as an adjustment variable for the company. To this extent it is not uncommon to see that employees with a long tenure are somewhat protected. For example, terminations fees for the company are required by law to increase with seniority in the company in many countries. This will not be further modeled in this study.

When focusing on voluntary turnover, the observed rates across demographics categories can be explained by the age pattern. Assume that the number of voluntary terminated employees of age  $a$  and tenure  $t$  is  $NVT(a, t)$  and that the number of employees of age  $a$  is  $NA(a)$  and the number of employee of tenure  $t$  is  $NT(t)$ , we get:

$$\text{Projected Tenure Pattern Voluntary Termination Rate}_a = \sum_{\text{Employee of age } a} \sum_{\text{Tenure } t} \frac{NVT(a, t)}{NA(a) \cdot NT(t)} \quad (7)$$

When comparing this projection (referred as "Tenure Pattern Projection") to the voluntary turnover rates presented in figure 8 (referred as "age pattern"), results are similar, as described in the figure 10.

An aged based model of the voluntary turnover rate  $\mu(a, t) = \mu^{voluntary}(a)$  can therefore be defined. The following exponential form is used:

$$\mu^{voluntary}(a) = C_{voluntary} \cdot e^{-\lambda_{age}^v \cdot a} \quad (8)$$

Coefficients	Value
$\lambda_{age}^v$ ( $year^{-1}$ )	0.062
$C_{voluntary}$ ( $year^{-1}$ )	0.6328
$R^2$	0.919

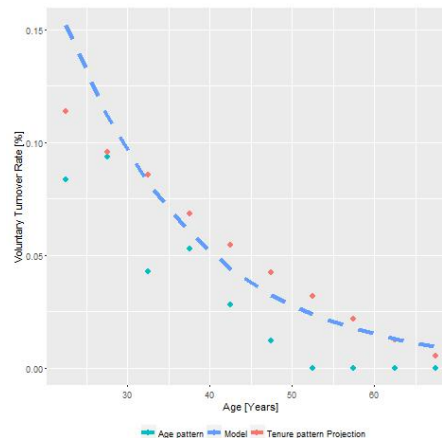


Figure 10: Termination Rate Mode

### 2.3 Salesforce Activity and Demographic Drivers:

The available data translated into the activity graphs of Figures 11, 12 and 13.

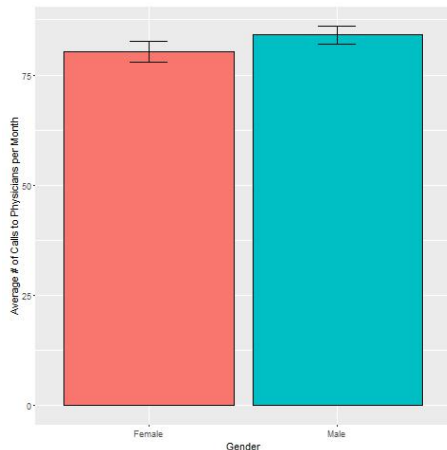


Figure 11: Physician Detailing Activity by Gender

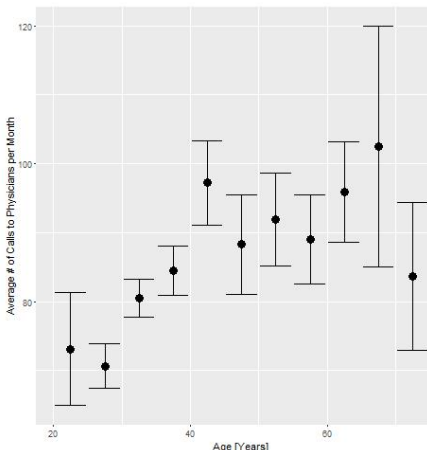


Figure 12: Physician Detailing Activity with Age

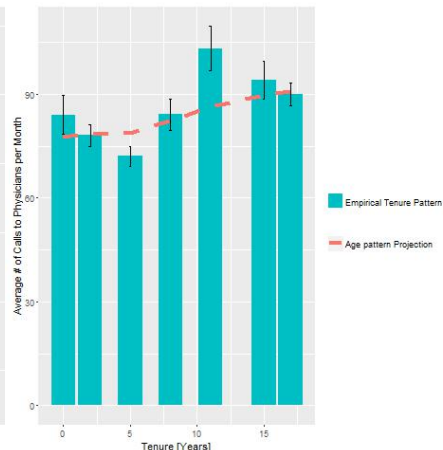


Figure 13: Physician Detailing Activity with Tenure

The overall activity pattern for a salesperson shows about 80 calls per month to a physician. No significant difference in activity among genders was observed. Yet the detailing activity seems to increase with age in a monotonic fashion (Spearman correlation = 0.8545). The activity pattern related to tenure seemed strange at first but finally seems to mainly result from the age composition of the tenure bands as shown on the graph above. The average detailing activity by age has been projected on tenure  $t$  categories in the following fashion. Assume that the detailing activity of employee of age  $a$  and tenure  $t$  at brick  $i$  is  $D_i(a, t)$  and that the number of employees of tenure  $t$  is  $NE(t)$ , we get:

$$\text{Detailing activity age projection}_t = \sum_{\text{Employee of tenure } t} \sum_{\text{Age } a} \frac{\mathbb{E}_i(D_i(a, t))}{NE(t)} \quad (9)$$

Note that 2 tenure bands (namely the 4 to 6 and 10 to 12 years ones) behaves in an unexpected fashion. According to this overall finding, the salesforce call activity  $D$  will only be considered as depending in age.

In a more general set-up, these findings are consistent with the concept of learning curve. This describes the human productivity improvement subsequent to task repetition (for a recent review, see [Anzanello and Fogliatto, 2011]). Their mathematical expression presents variations around 3 main forms: log linear, exponential and hyperbolic. These variations are justified by the introduction of notions such as previous experience, cost or production time threshold or learning heterogeneity among workers. Note that some research has also focused on the topic of learning and forgetting cycle [Jaber and Bonney, 1997] as well as the impact of stochastic work conditions [Qin and Nembhard, 2010]. Nonetheless these features are out of scope of this study.

Among the canonical learning curve expressions, a 2015 meta-analysis [Grosse et al., 2015] suggested that the best fits were obtained through hyperbolic or exponential schemes. They will therefore be the ones used in this paper.

The following form will be used to describe the activity uptake with age:

$$D(a) = k \cdot \frac{a}{a + R} \quad (10)$$

Non linear least square calibration based on the data used in figure 12 yields the following results:

Coefficient	Value	Std
k	112.039	9.224
R	12.384	4.606

Model intercept (red line) has been reported on the graph on the right.

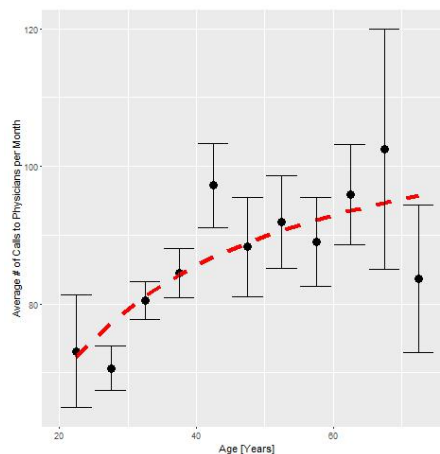


Figure 14: Modeled Detailing Activity with Age

## 2.4 Promotion Response and Demographic Drivers:

Now that we have seen that demographic factors have an impact on turnover and on call volumes, we will now look at the impact of salesforce demographics in the sales dependency in call volume parameter  $\beta$  used in in equation (1). This will be done by adapting the marketing mix framework built in as equation (1) described in the appendix. Repeating the calibration procedure on the data used in the previous section yields the results shown in Figures 15, 16 and 17.

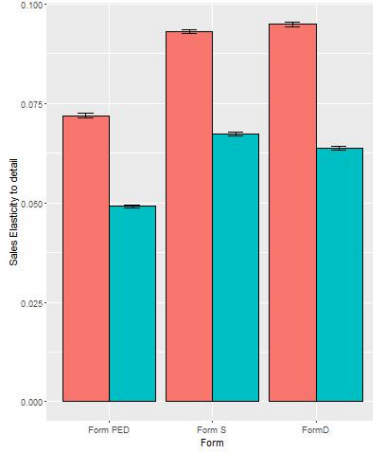


Figure 15: Sales Sensitivity to Detailing  $\beta$  by Gender

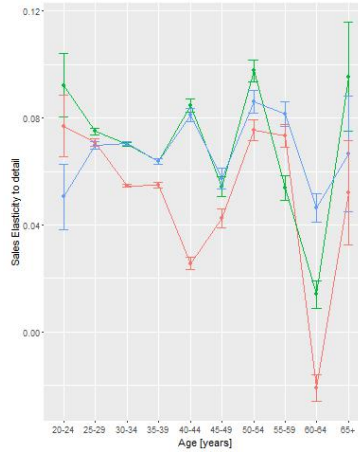


Figure 16: Sales Sensitivity to Detailing  $\beta$  with Age

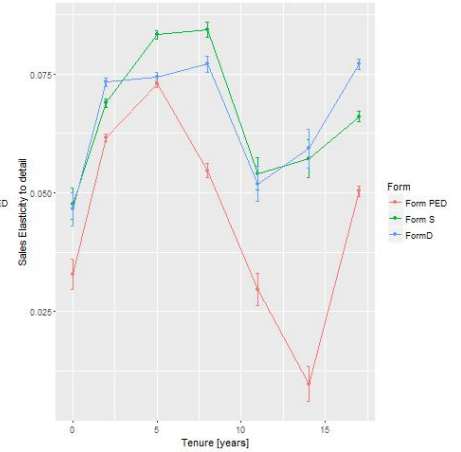


Figure 17: Sales Sensitivity to Detailing  $\beta$  with Tenure

The key outcome here is a clear difference in detailing efficiency across genders. This difference is consistent for the 3 products forms and suggest that female salesperson activity drives more sales than their male counterpart. Gender yet needs to be considered with extreme caution. From a legal standpoint discrimination is strictly forbidden. As such the quantitative argument presented above can only be used to promote the establishment of a fully balanced salesforce with 50% female and 50% male from the current 43% female.

Tenures and age patterns seem more difficult to explain at first sight. But they result from the gender variations inside of age and tenure bands. This can indeed be infered because of the linear structure of the marketing mix model at brick level and because calibration shows no difference in the main parameters and  $R^2$  over the 3 modified frameworks.

Detailing efficiency is therefore considered as being only dependent in gender in the rest of this study. Note that in the rest of the paper, we will not further discuss gender balance (i.e the % of female in the workforce) and assume that the current state of 43% of female is maintained.

### 3 How to Create a comprehensive sales force Framework?

To our knowledge if the mix of PDE and learning curve has been investigated [Gerchak et al., 1990] and the impact of learning and turnover has been investigated in some fashion [Gans and Zhou, 2002], there has been no attempt to integrate the notions of turnover and learning in a continuous PDE framework. A possible reason might be the difficulty to converge to simple closed formulas and the necessity to run complex optimization programs as stressed in [Hewitt et al., 2015] .

#### 3.1 Sales Force Activity ROI in a demographic framework:

**Valuation Framework.** To embed voluntary turnover in the marketing mix framework, a Markov chain  $X_{t,i}$  representing the availability of the salesperson at time  $t$  on brick  $i$  is introduced.  $X_{t,i}$  can either equal 1 when the salesperson is in the field promoting the product or 0 when the salesperson

has left the company.

$$\begin{aligned} P(X_{t+1} = 0|X_t = 1) &= dt \cdot \mu(a) \\ P(X_{t+1} = 1|X_t = 0) &= 1. \end{aligned} \quad (11)$$

Note that because of monthly granularity of the analysis, the turnover probability has been rescaled by a factor  $dt = \frac{1}{12}$ . Also note that vacancy has only been supposed to last for a month with this current setup: if an employee has left at a given month  $t$ , she/he is replaced at the month  $t + 1$  with probability 1. Cumulating turnover consideration with gender ones from the previous section then leads to the following revised marketing mix framework (adjusted from equation 1):

$$\begin{aligned} \log(\text{sales}_{i,t,f} + 1) &= \sum_{q=1}^{q=4} Q_{q,f} + \sum_{o=1}^{o=0} \log(\text{sales}_{i,t-o,f} + 1)\theta_{o,f} + \sum_{\text{gender } g} \beta_{f,g} \log(D_{i,t,g} \cdot X_{t,i} + 1) + \\ &\lambda_f^{\text{brand}} \Delta \log(p_{i,t,f}^{\text{brand}} + 1) + \lambda_f^{\text{Competition}} \Delta \log(p_{i,t,f}^{\text{Competition}} + 1) + \epsilon_{i,t,f} \end{aligned} \quad (12)$$

Finally accounting for call productivity among age bands is done by correcting the  $C_d$  term used in Section 1 (defined as average the call unit cost, see Section 1.3) by the harmonic function used to described the activity dependency in age:

$$C_d(a) = C_d \cdot \frac{a + R}{k \cdot a} \quad (13)$$

Assuming all workforce members bear the same full costs (base and variable salary plus taxes and benefits) for a job, the cost a call is approximated by the cost of worker divided by its average number of call on a monthly basis defined in (10). So that, according to the notation used in 1, the average ROI of the call activity  $D$  at brick level  $i$  (initially defined in 5) becomes a function of age  $a$  and gender  $g$  over the forms  $f$ :

$$\Pi(P_f, D, a) = \mathbb{E}_i(\mathbb{E}_X(\sum_{t=0}^{t=36} P_{f,i} \cdot \text{sales}_{i,t,f}(P_{f,i,f}, D_i, X_t))) \quad (14)$$

$$ROI_{\text{Call Activity}}^g(D, a) = \frac{\sum_f \Pi(P_f, D, a) - \Pi(P_f, 0, a)}{T \cdot C_d(a) \cdot D}. \quad (15)$$

**Empirical Results.** The parameters calibrated in the two previous section were used to simulate the  $ROI_{\text{Call Activity}}^g(D, a)$  function according to the equation 15. Assuming, the same level of call activity  $D$  stays constant, two ROIs functions of age were obtained (one per gender). Averaging the results led to the graph below in Figure 18.

The ROI structure displayed on the right was estimated by keeping the call volume constant and changing the associated salesperson profile. The result shows a structure that is again in line with the notion of learning curve. More experienced sales representative yield an increase of ROI by about 100%. If the change of ROI among employee categories is interesting the direct action in this study case should be to cut the call activity for the product and to reallocate the salesforce. Also note that compared a non demographic framework this enhancement shows significant differences because of turnover adjustments ( $\Delta$  in ROI > 50%).

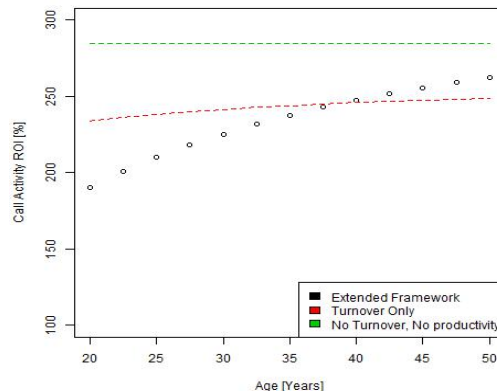


Figure 18: Average ROI - 1Y - Simulations

This ROI structure can be tied back again to a learning curve concept. Salesperson become more efficient over time because of an increased stability and efficiency in the job. The call activity ROI described in 18 has a complex analytical form (cf 15). This form can be approximated. This approximation will be referred as  $\omega$  over the following paragraph and will take the following form as:

$$\omega(a) = C_{\omega} \cdot a^{b_{\omega}} \quad (16)$$

Calibrating the form in equation 16 with the data used to build 18 resulted in a  $R^2$  of 99% and the following coefficients:

Coefficient	Value	Std
$C_{\omega}$	4.225981	0.035084
$b_{\omega}$	0.347302	0.009943

### 3.2 An age and tenure structured model to describe the workforce

The sales representatives workforce is described in a structured equations framework. Its density will be noted  $\rho(t, a, z)$  and depends in the time  $t$ , in the age  $a$  and in the tenure in company  $z$ . These 3 variables help frame the problem. Time connects the workforce to business events (in this case the promoted product lifecycle), age (bounded between  $a_{min}$  and  $a_{max}$ ) is used as a proxy to experience to determine sales representatives profile and finally tenure in the company is used as an input for the learning curve. The framework focuses on two main workforce dynamics, namely turnover and hiring. Turnover is represented by  $\mu$ , the departure rate of workforce at a given time  $t$  according to age and tenure. Hiring on the other hand is handled as a boundary condition at tenure  $z = 0$  and is described by two quantities: the hiring rate  $h_t$ , which represents the percentage of new hires over the total workforce  $P_t$  at time  $t$  and the hiring density  $\gamma(a)$ , which represents the hiring profile repartition across age categories. This can be summarized through the following set of equations:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial a} + \frac{\partial \rho}{\partial z} = -\mu \cdot \rho \quad (17)$$

$$\rho(t, a, 0) = h_t \cdot P_t \cdot \gamma(a) \quad (18)$$

with:

$$P_t = \int_{a_{min}}^{a_{max}} \int_0^{a-a_{min}} \rho(t, a, z) dz da \quad (19)$$

Workforce ROI (resp. costs) is addressed through  $\omega(a, z)$  (resp.  $c(a, z)$ ) and represents the output in sales associated to one unit of workforce referred as a headcount (resp. the full labor costs borne by the firm for one headcount). These two functions are assumed to be age and tenure dependent.

### 3.3 Sales force Demographic Framework and Steady State hiring criterion

**Demographic Framework.** Based on the previous considerations, it becomes easy to create a workforce with the right size and demographic composition. Assume  $\rho(a, z)$  to be the number of employee of age  $a$  (bounded between  $a_{min}$  and  $a_{max}$ ) and tenure  $z$  and that (according to the previous section), the turnover rate  $\mu$  is only age dependent. The hiring strategy  $H(a)$  has to be nonnegative and determines the equilibrium headcount through the steady state of the following equation:

$$\begin{cases} \frac{\partial \rho^*}{\partial a} + \frac{\partial \rho^*}{\partial z} = -\mu(a)\rho^*, \\ \rho^*(a, 0) = H(a) \geq 0. \end{cases} \quad (20)$$

Assume we want to show the value of changing a salesforce composition, we will assume that the call activity  $D$  is predefined. This leads to:

$$D = \int_{a_{min}}^{a_{max}} \int_0^{a-a_{min}} \rho^*(a, z) D(a) dz da. \quad (21)$$

While the main objective is to maximize the overall salesforce ROI:

$$ROI([H]) = \int_{a_{min}}^{a_{max}} \int_0^{a-a_{min}} \rho^*(a, z) \omega(a) dz. \quad (22)$$

This leads to an optimal hiring strategy at an optimal age  $a^*$  which depends on the call activity  $D$  and on the call ROI  $\omega$  in a complex manner.

First, there is a need to identify the analytical expression of  $\rho^*$  for a given hiring profile  $H(a)$ . This is determined by the solution of (20) thanks to the method of characteristics :

$$\rho^*(a, z) = H(a - z) e^{-\int_0^z \mu(a - z + \sigma) d\sigma}.$$

Then the overall call activity constraint can be written as:

$$\begin{aligned} D &= \int_{a_{min}}^{a_{max}} \int_0^{a-a_{min}} H(a - z) e^{-\int_0^z \mu(a - z + \sigma) d\sigma} D(a) dz da \\ &= \int_{a_{min}}^{a_{max}} \int_0^{a_{max}-b} H(b) e^{-\int_0^z \mu(b + \sigma) d\sigma} D(b + z) dz db, \end{aligned}$$

and thus

$$D = \int_{a_{min}}^{a_{max}} H(b) f(b) db, \quad f(b) = \int_0^{a_{max}-b} e^{-\int_0^z \mu(b + \sigma) d\sigma} D(b + z) dz.$$



The same calculation allows us to write

$$ROI(H) = \int_{a_{min}}^{a_{max}} H(b)g(b)db, \quad g(b) = \int_0^{a_{max}-b} \frac{1}{e} \int_0^z \mu(b+\sigma)d\sigma \omega(b+z)dz.$$

To maximize this expression, we just observe that, because  $H \geq 0$

$$ROI([H]) \leq \max_u \frac{g(u)}{f(u)} \int_{a_{min}}^{a_{max}} f(b)H(b)db = D \max_u \frac{g(u)}{f(u)},$$

and the equality is reached for

$$H(b) = h \delta(b - a^*), \quad \frac{g(a^*)}{f(a^*)} = \max_u \frac{g(u)}{f(u)}, \quad (23)$$

and the total hiring rate is determined by the constraint  $D$ , which gives

$$h = \frac{D}{f(a^*)}.$$

**Empirical Results.** Hence, using the formulas (10) for the voluntary turnover rate, (12) describing the activity uptake with age detailing efficiency and (18) for the call activity ROI, simulations of the model (17)–(19) are easily computed to find the optimal age and hiring profile to maximise the overall salesforce ROI.

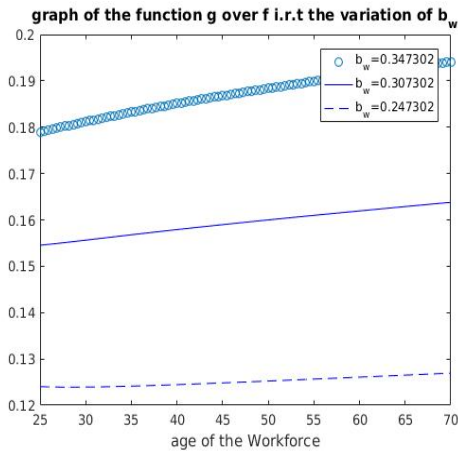


Figure 19: Graph of the function  $g$  over  $f$  to maximize with the variation of parameter  $b_w$

Let us assume that the minimum (resp. maximum) age of an employee  $a_{min}$  (resp.  $a_{max}$ ) is 25 (resp. 70) while the overall predefined call activity  $D$  is at 200.000. The framework suggests (figure 19) an optimal hiring age of 70 years. If this result can be expected because of the structure of the call activity  $D$  and its ROI  $\omega$ , it remains highly unrealistic as a sourcing strategy. It is therefore possible to get the right size and demographic composition given a predefined call activity. The continuous model leveraging the approximated ROI form 16 leads to the simple closed formula when optimized :

$$\rho^*(a, z) = \rho^*(a, a-a^*) = \frac{D}{f(a^*)} e^{-\int_0^{a-a^*} \mu(a^*+\sigma)d\sigma}$$

The optimal age  $a^*$  in the expression above is either the minimum or the maximum age of an employee. This appears an an evidence with the previous sections. The call activity uptake (small value of  $R$ ) with age is the main phenomenon at stake, its appears much more interesting to capitalize on experienced employees.

**Demographic framework adding tenure on the call activity ROI** Let us now explore what happens when tenure considerations are added to the demographic framework. For instance, let us assume that the call activity ROI of a sales person  $\omega$  can be estimated by :

$$\omega(a, z) = C_{\omega} \cdot (a + z)^{b_{\omega}}. \quad (24)$$

This implies that the overall salesforce ROI to maximize is now determined by :

$$ROI([H]) = \int_{a_{min}}^{a_{max}} \int_0^{a-a_{min}} \rho^*(a, z) \omega(a, z) dz. \quad (25)$$

This change now induces a shift in behaviour of the function  $g$ , which can be written as :

$$g(b) = \int_0^{a_{max}-b} e^{-\int_0^z \mu(b+\sigma) d\sigma} \omega(b+z, z) dz.$$

Including these modifications, the optimization problem leads to :

$$H(b) = h \delta(b - a^*), \quad \frac{g(a^*)}{f(a^*)} = \max_u \frac{g(u)}{f(u)}, \quad (26)$$

and the total hiring rate is determined by the constraint  $D$ , which gives

$$h = \frac{D}{f(a^*)}.$$

With this new model, the optimal age and tenure obtained in order to maximize the overall salesforce ROI strongly depends on the parameters of the activity uptake and the call activity ROI.

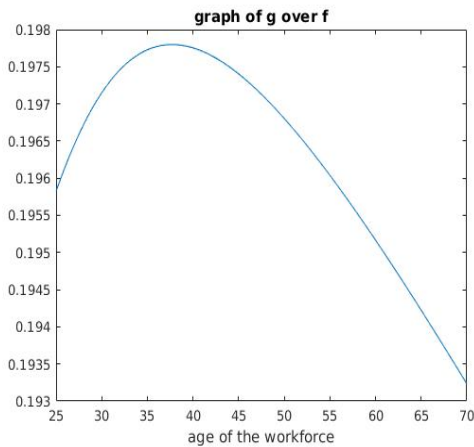


Figure 20: Graph of the function  $g$  over  $f$  to maximize with  $b_{\omega} = 0.347302$  and  $R = 12$

Compared to the previous framework, the optimal solution change drastically. The hiring policies that maximizes the salesforce ROI sets a hiring age of  $a^* = 37.7$  years (under the assumptions that  $R = 12$  and  $b_{\omega} = 0.347302$ ). Moreover it can be shown that the more the parameter  $R$  grows (*i.e.* the less difference in call productivity there is between young and experienced representatives), the more the hiring policy would recommend focusing on young people. Adding tenure consideration to the mix therefore changes drastically the possible implications of the model. As such, one should pay attention to the variables used in the model.

### 3.4 Discussion.

The PDE framework and its integration with the marketing mix one leads to interesting consideration in terms of hiring profile and sales force sizing. The results displayed in this section indeed show that one can optimize its salesforce composition and size while keeping its call activity constant to avoid too much disruption. We also argue that the optimization can be further pursued to optimize in the same time both the salesforce composition and the call activity to maximize the salesforce effectiveness. Yet several limitations appears. The choices made in the modeling seem to highly influence the results, notably in terms of the demographic variable selected for the analysis. Relying only on age seems to create unrealistic scenarios while integrating tenure consideration may provide a more balanced picture.

Several constraints associated to the implementation of such results also appear. First this kind of study makes little sense for sales force with an overall call activity ROI below or close to 100%. Optimizing the ROI is therefore a priority, then optimizing the sourcing strategy can yield additional results. This type of approach would yield a lot of sense for key account management type of position. Second the age approach may not be suitable for all labor market. For example this would not be possible in geographies such as the UK or the US. Other constraints such as legal retirement age and specific compensatoin packages are not included here.

Also note that the ROI valorisation with turnover was based on a one-year time horizon. Additional computational ideas would be required to efficiently expand this framework to a 3 to 5 year framework. Finally sales force short term dynamics and optimization should be better discussed. The optimization in the last section is indeed based on steady state hypothesis that can be challenged with the short term valuation framework that is being used. It would prove more interesting and realistic to consider a dynamical approach of the problem. The continuous model in the last section and the marketing mix framework in the previous one are inherently linked. It could therefore be important to assess the magnitude of the marketing mix model oscillations associated to the recommended workforce shifts.

## Conclusion and Next Steps

**Conclusion.** Standard regressions techniques used in marketing research are instrumental in providing robust analytical insights for sales force design and optimization. They also have a strong explanatory power regarding strategic marketing decisions linked to product pricing and promotional mix allocation. Yet summarizing the salesforce to its detailing activity can be misleading. It can be shown that demographic parameters such as gender, age and tenure yield important differences on turnover, productivity and detailing efficiency. Though it is difficult to draw standard concepts, because of contexts that can heavily change among salesforce and product, integrating demographics into a comprehensive salesforce valuation framework helps perfoming informed salesforce investment decision. As shown in this paper, turnover can represent important market disruption and has therefore to be integrated to discount the apparent ROI of the salesforce associated to its workload. This is completely manageable when Human Resources and Marketing intelligence work together, and shows promising results in light of this study.

**Next Steps.** Several next steps have been planed. In terms of pure research, questions around the drivers of turnover have emerged. High salesforce turnover and therefore high ROI impacts are common. A natural next step is therefore to propose a more detailed framework to design mitigation actions. A second area of interest (research-wise) lies in a better integration in such frameworks of real

world evidence type of insights related to product therapeutic value. To our knowledge, this has not yet been done. Standard marketing techniques indeed rely on finding analog products and replicating their characteristics rather than digging into their core patient characteristics.

Finally a question of automation and insight industrialization arises. In a current setup, information along with its contextual knowledge is heavily disseminated. Strong transversal collaborations are therefore required (for instance human resources and marketing) to create value added insights.

**Acknowledgments.** Marie Doumic and Mathieu Mezache were supported by the ERC Starting Grant SKIPPER<sup>AD</sup> (number 306321). Benoit Perthame and Delphine Salort were supported by the French "ANR blanche" project Kibord: ANR-13-BS01-0004 funded by the French Ministry of Research.

## References

- [Andris A. Zoltners, 2006] Andris A. Zoltners, Prabhakant Shinha, S. E. L. (2006). Match your sales force structure to your business life cycle. *Harvard Business Review*.
- [Anzanello and Fogliatto, 2011] Anzanello, M. J. and Fogliatto, F. S. (2011). Learning curve models and applications: Literature review and research directions. *International Journal of Industrial Ergonomics*, 41(5):573 – 583.
- [Calfee et al., 2002] Calfee, J., Winston, C., and Stempski, R. (2002). Direct-to-consumer advertising and the demand for cholesterol-reducing drugs. *Journal of Law and Economics*, 45(S2):673–690.
- [Cook, 2015] Cook, M. A. G. (2015). *Forecasting for the pharmaceutical industry: models for new product and in-market forecasting and how to use them*. Gower Publishing, Ltd.
- [Darmon, 1982] Darmon, R. Y. (1982). Identifying profit-producing salesforce members. *Journal of Personal Selling & Sales Management*, 2(2):14–23.
- [Darmon, 1990] Darmon, R. Y. (1990). Identifying sources of turnover costs: A segmental approach. *Journal of Marketing*, 54(2):46–56.
- [Darmon, 2004] Darmon, R. Y. (2004). Controlling sales force turnover costs through optimal recruiting and training policies. *European Journal of Operational Research*, 154(1):291–303.
- [Desiraju et al., 2004] Desiraju, R., Nair, H., and Chintagunta, P. (2004). Diffusion of new pharmaceutical drugs in developing and developed nations. *International Journal of Research in Marketing*, 21(4):341 – 357. Special Issue on Global Marketing.
- [Domic et al., 2016] Doumic, M., Perthame, B., Ribes, E., Salort, D., and Toubiana, N. (2016). Toward an integrated workforce planning framework using structured equations. *ArXiv e-prints*.
- [Gans and Zhou, 2002] Gans, N. and Zhou, Y.-P. (2002). Managing learning and turnover in employee staffing. *Operations Research*, 50(6):991–1006.
- [Gerchak et al., 1990] Gerchak, Y., Parlar, M., and Sengupta, S. S. (1990). On manpower planning in the presence of learning. *Engineering Costs and Production Economics*, 20(3):295–303.

- [Grosse et al., 2015] Grosse, E. H., Glock, C. H., and Mller, S. (2015). Production economics and the learning curve: A meta-analysis. *International Journal of Production Economics*, 170, Part B:401 – 412. Current Research Issues in Production Economics.
- [Hahn et al., 1994] Hahn, M., Park, S., Krishnamurthi, L., and Zoltners, A. (1994). Analysis of new product diffusion using a four segment trial-repeat model. *Marketing Science*, 13:224–247.
- [Hewitt et al., 2015] Hewitt, M., Chacosky, A., Grasman, S. E., and Thomas, B. W. (2015). Integer programming techniques for solving non-linear workforce planning models with learning. *European Journal of Operational Research*, 242(3):942 – 950.
- [Jaber and Bonney, 1997] Jaber, M. Y. and Bonney, M. (1997). A comparative study of learning curves with forgetting. *Applied Mathematical Modelling*, 21(8):523 – 531.
- [Karen E. Flaherty, 2002] Karen E. Flaherty, J. M. P. (2002). Using career stage theory to predict turnover intentions among salespeople. *Journal of Marketing Theory and Practice*, 10(3):48–57.
- [Kremer et al., 2008] Kremer, S. T., Bijmolt, T. H., Leeflang, P. S., and Wieringa, J. E. (2008). Generalizations on the effectiveness of pharmaceutical promotional expenditures. *International Journal of Research in Marketing*, 25(4):234 – 246. Marketing and Health.
- [Leonard Jon Parsons, 1981] Leonard Jon Parsons, P. V. A. (1981). *Analysis of Sales Call Effectiveness*, volume 18. American Marketing Association.
- [Lilien et al., 1981] Lilien, G. L., Rao, A. G., and Kalish, S. (1981). Bayesian estimation and control of detailing effort in a repeat purchase diffusion environment. *Management Science*, 27(5):493–506.
- [Lucas et al., 1987] Lucas, G. H., Parasuraman, A., Davis, R. A., and Enis, B. M. (1987). An empirical study of salesforce turnover. *Journal of Marketing*, 51(3):34–59.
- [Manchanda and Chintagunta, 2004] Manchanda, P. and Chintagunta, P. (2004). Responsiveness of physician prescription behavior to salesforce effort: An individual level analysis. *Marketing Letters*, 15(2-3):129–145.
- [Muchinsky and Morrow, 1980] Muchinsky, P. M. and Morrow, P. C. (1980). A multidisciplinary model of voluntary employee turnover. *Journal of Vocational Behavior*, 17(3):263 – 290.
- [Ng and Feldman, 2009] Ng, T. W. and Feldman, D. C. (2009). Re-examining the relationship between age and voluntary turnover. *Journal of Vocational Behavior*, 74(3):283 – 294.
- [Peres et al., 2010] Peres, R., Muller, E., and Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, 27(2):91 – 106.
- [Qin and Nembhard, 2010] Qin, R. and Nembhard, D. A. (2010). Workforce agility for stochastically diffused conditions—a real options perspective. *International Journal of Production Economics*, 125(2):324–334.
- [Stremersch, 2008] Stremersch, S. (2008). Health and marketing: The emergence of a new field of research. *International Journal of Research in Marketing*, 25(4):229 – 233. Marketing and Health.

- [Vakratsas and Kolarici, 2008] Vakratsas, D. and Kolarici, C. (2008). A dual-market diffusion model for a new prescription pharmaceutical. *International Journal of Research in Marketing*, 25(4):282 – 293. Marketing and Health.
- [Wright, 1936] Wright, T. P. (1936). Factors Affecting the Cost of Airplanes. *Journal of Aeronautical Sciences*, 3:122–128.

## A Appendix - Demographics and PRA - Details.

**Gender.** Reviewing the previous framework (cf equation 1) to allow differentiation in detailing impact across gender yields the following framework:

$$\log(\text{sales}_{i,t,f} + 1) = \sum_{q=1}^{q=4} Q_{q,f} + \sum_{o=1}^{o=O} \log(\text{sales}_{i,t-o,f} + 1)\theta_{i,o,f} + \sum_{g=\text{gender}} \beta_{i,f,g} \log(D_{i,g,t} + 1) + \lambda_{i,f}^{\text{brand}} \Delta \log(p_{i,t,f}^{\text{brand}} + 1) + \lambda_{i,f}^{\text{Competition}} \Delta \log(p_{i,t,f}^{\text{Competition}} + 1) + \epsilon_{i,t,f} \quad (27)$$

A calibration similar to the one employed in the first section of this paper translates into the following results.

Coefficient	Form A	Form B	Form C
$Q_1$	-	6.612721 (0.223701)	6.735391 (0.113897)
$Q_2$	-	6.692029 (0.228606)	7.064003 (0.116864)
$Q_3$	-	6.784395 (0.223276)	7.282193 (0.115043)
$Q_4$	-	6.511399 (0.222607)	6.651347 (0.113290)
$\theta_1$	0.650298 (0.002095)	0.655954 (0.002047)	0.456340 (0.003333)
$\theta_2$	-	-	0.086706 (0.003736)
$\theta_3$	-	-	0.172991 (0.003424)
$\beta_{\text{male}}$	0.067221 (0.002877)	0.063752 (0.002996)	0.049073 (0.002736)
$\beta_{\text{female}}$	0.092954 (0.003201)	0.094840 (0.003335)	0.071915 (0.003044)
$\lambda^{\text{brand}}$	-0.068788 (0.016968)	-1.171467 (0.070917)	-2.474504 (0.043196)
$\lambda^{\text{competition}}$	0.452823 (0.023089)	-0.422875 (0.015001)	0.664828 (0.024732)
$R^2$	0.9347	0.9399	0.9511
$\sigma_{t,f}$	0.7102	0.7435	0.6922

**Age.** Reviewing the previous framework to allow differentiation in detailing impact across age bands yields the following framework:

$$\log(\text{sales}_{i,t,f} + 1) = \sum_{q=1}^{q=4} Q_{q,f} + \sum_{o=1}^{o=O} \log(\text{sales}_{i,t-o,f} + 1)\theta_{i,o,f} + \sum_{a=\text{age}} \beta_{i,f,a} \log(D_{i,a,t} + 1) + \lambda_{i,f}^{\text{brand}} \Delta \log(p_{i,t,f}^{\text{brand}} + 1) + \lambda_{i,f}^{\text{Competition}} \Delta \log(p_{i,t,f}^{\text{Competition}} + 1) + \epsilon_{i,t,f} \quad (28)$$

This translates into the following results.

Coefficient	Form A	Form B	Form C
$Q_1$	-	6.522338 (0.224788)	6.749270 (0.114568)
$Q_2$	-	6.604355 (0.229733)	7.079908 (0.117537)
$Q_3$	-	6.696412 (0.224395)	7.294906 (0.115717)
$Q_4$	-	6.420754 (0.223701)	6.664689 (0.113964)
$\theta_1$	0.650226 (0.002095)	0.655926 (0.002050)	0.455288 (0.003332)
$\theta_2$	-	-	0.086877 (0.003733)
$\theta_3$	-	-	0.173778 (0.003423)
$\beta_{20}$	0.092222 (0.010595)	0.050542 (0.010857)	0.077022 (0.010146)
$\beta_{25}$	0.074903 (0.004679)	0.069806 (0.004856)	0.070863 (0.004442)
$\beta_{30}$	0.070114 (0.003581)	0.070406 (0.003701)	0.054635 (0.003416)
$\beta_{35}$	0.063712 (0.004365)	0.063636 (0.004523)	0.054910 (0.004164)
$\beta_{40}$	0.084634 (0.006416)	0.080996 (0.006598)	0.025565 (0.006066)
$\beta_{45}$	0.054264 (0.007482)	0.057343 (0.007774)	0.042384 (0.007140)
$\beta_{50}$	0.097700 (0.006894)	0.086170 (0.007187)	0.075440 (0.006607)
$\beta_{55}$	0.053697 (0.008388)	0.081406 (0.008675)	0.073423 (0.007919)
$\beta_{60}$	0.013990 (0.008367)	0.046374 (0.008602)	-0.021159 (0.008010)
$\beta_{65}$	0.095392 (0.014753)	0.066481 (0.015614)	0.052053 (0.014139)
$\beta_{70+}$	0.11579 (0.015821)	0.106916 (0.016663)	0.078244 (0.015677)
$\lambda^{brand}$	-0.053399 (0.017122)	-1.173142 (0.071390)	-2.493668 (0.043586)
$\lambda^{competition}$	0.433232 (0.023298)	-0.398885 (0.015070)	0.685115 (0.025262)
$R^2$	0.9347	0.9399	0.9512
$\sigma_{t,f}$	0.71	0.7436	0.6917

**Tenure.** Reviewing the previous framework to allow differentiation in detailing impact across tenure bands yields the following framework:

$$\begin{aligned}
\log(\text{sales}_{i,t,f} + 1) = & \sum_{q=1}^{q=4} Q_{q,f} + \sum_{o=1}^{o=O} \log(\text{sales}_{i,t-o,f} + 1)\theta_{i,o,f} + \sum_{z=\text{tenure}} \beta_{i,f,z} \log(D_{i,z,t} + 1) + \\
& \lambda_{i,f}^{brand} \Delta \log(p_{i,t,f}^{brand} + 1) + \lambda_{i,f}^{Competition} \Delta \log(p_{i,t,f}^{Competition} + 1) + \epsilon_{i,t,f}
\end{aligned} \tag{29}$$

This translates into the following results.



Coefficient	Form A	Form B	Form C
$Q_1$	-	6.493609 (0.224324)	6.700192 (0.114175)
$Q_2$	-	6.576578 (0.229230)	7.030938 (0.117104)
$Q_3$	-	6.668710 (0.223904)	7.248457 (0.115287)
$Q_4$	-	6.391826 (0.223233)	6.616312 (0.113556)
$\theta_1$	0.650992 (0.002094)	0.656081 (0.002048)	0.455642 (0.003334)
$\theta_2$	-	-	0.086815 (0.003735)
$\theta_3$	-	-	0.173700 (0.003425)
$\beta_0$	0.047681 (0.005894)	0.046563 (0.006103)	0.032806 (0.005614)
$\beta_2$	0.068884 (0.004140)	0.073275 (0.004287)	0.061497 (0.003949)
$\beta_5$	0.083271 (0.003998)	0.074285 (0.004126)	0.072910 (0.003805)
$\beta_8$	0.084286 (0.005346)	0.077025 (0.005528)	0.054637 (0.005107)
$\beta_{11}$	0.053955 (0.007489)	0.051882 (0.007792)	0.029672 (0.007127)
$\beta_{15}$	0.057204 (0.006726)	0.059275 (0.006920)	0.009797 (0.006326)
$\beta_{17+}$	0.065955 (0.004064)	0.077038 (0.004221)	0.050315 (0.003900)
$\lambda^{brand}$	-0.064073 (0.017100)	-1.161241 (0.071101)	-2.463595 (0.043284)
$\lambda^{competition}$	0.447361 (0.023265)	-0.401935 (0.014910)	0.666625 (0.024901)
$R^2$	0.9347	0.9399	0.9511
$\sigma_{t,f}$	0.7104	0.7437	0.6921