Abstract: This paper investigates the impact of socio-economic factors on the unemployment spell of fresh graduates. The data sample is drawn from the Living Standards Measurement Survey of Albania in 2012. We restrict the sample to only those aged from 19 to 35 years old to reduce the heterogeneity among unemployed graduates. In addition, we drop all observations with unemployment spell greater than 19 years given that the minimum working age in Albania is 16. The job-search theory constitutes the theoretical framework of this work. The socio-economic factors we examine in this paper are age, gender, education attainment, residence and occupation. The empirical part employs the Gompertz hazard proportional model to determine the effect of socio-economic factors on the first unemployment period of fresh graduates. Our findings indicate that the Gompertz model provides a better fit to the data based on the AIC value compared to other models including the log-logistic and Cox models. Further, the estimates in the Gompertz specification are in line with the theoretical findings of the job-search theory. Education, age, gender and occupation are significant determinants of the first unemployment spell. In all models, the indicator of urban residence results insignificant. Moreover, unemployment duration increases with age. Lastly, those with high levels of education and high occupation profiles have higher odds of exiting from unemployment.

Key words: unemployment spell, hazard rate, proportional hazard, Gompertz

1. INTRODUCTION

A n increasing number of publications in transition analysis have paid special attention to school to work transition. Most of the studies suggest that graduates with less favorable socio-economic background are expected to have longer unemployment spells. Furthermore, schooling and the characteristics of the region where the unemployed graduate comes from play an important role in the probability of finding a job. For instance, Lassibille et al (2001) examine the impact of human capital on unemployment duration using a data set drawn from the Encuesta Socio-Demográfica in Spain in 1991. They employ multinomial logistic models and find that males and those with high education profiles have shorter unemployment spells compared to females and those with low school attainment. In addition, their results indicate insignificance of family background on the probability of exiting from unemployment. Jaunky and Khadaroo (2007) study the unemployment duration of graduates from universities in Mauritania from 1995 to 2000 using several proportional hazard and accelerated failure time models. The main findings suggest that age affects negatively the first duration of unemployment. Further, the father's education is positively correlated with the graduate's unemployment spell. Canals and Diebolt (2002) employs a survival model to determine the effect of education background on the unemployment duration of French graduates in 1996. The authors find that the average spell of tertiary education graduates is roughly 4 months. In addition, the mean duration of graduates from Business schools is
approximately 2 months, and those graduated in Law, Economics and Social Sciences are more likely to experience longer spells. In Albania, Alikaj and Shehaj (2015) use Living Standards Measurement Survey (LSMS, 2012) to investigate the determinants of school to work transition. The authors estimate a Weibull model and conclude that the unemployment duration of those employed in high profile jobs (e.g., professionals and technicians) have higher probabilities of exiting from unemployment.

This paper builds on Gorenca (2018) which provides the theoretical framework of the job-search theory. In this study, we omit the theoretical model, which explicitly derives the determinants of unemployment duration (Mortensen, 1986). However, section (2) provides a summary of the expected effect of the main determinants on unemployment spells. Furthermore, we borrow from parametric estimations to examine the effect of factors including age, gender, education attainment, residence and occupation on unemployment duration. The data sample is drawn from the LSMS (2012). The main data limitation is the absence of the reservation wage, a measure of unemployment insurance and social assistance from the dataset.

This paper is organized as follows. Section (2) provides an overview of the labor market situation in Albania. Section (3) presents the building blocks of duration analysis and the Gompertz model, and section (4) presents the data and estimation results. Lastly, section (5) concludes.

2. DETERMINANTS OF UNEMPLOYMENT DURATION

This section identifies the main factors that determine the employment probability of jobseekers. In essence, employment determinants are derived from the job search theory (Gorenca, 2018). Mortensen (1977) argues that the probability of exiting unemployment is the probability of receiving a new job offer times the probability of accepting it. Collier (2003) claims that individual characteristics do also play a critical role in unemployment duration as they determine the reservation wage and individual labor preferences. In this work, we concentrate on individual characteristics including age and gender. Regarding human capital attributes, the literature survey highlights the impact of labor market experience, education attainment, and training qualifications (Borjas, 2010; Cahuc & Zylberberg, 2004; Berndt, 1991). However, in this paper we neglect the effect of experience owing to the heterogeneity of jobseekers, i.e., not all of the jobseekers were previously employed. Other determinants examined by Gorenca (2018) are unemployment benefits and social insurance.

The job search theory anticipates that under certain circumstances, the high tenure of receiving unemployment payments will lead to longer unemployment spells (Bover, Arellano & Bentolila, 2001; Mortensen, 1977; Pellizari, 2006). Borjas (2010) claims that there must be a positive relationship between the economic replacement rate and unemployment duration. For instance, Moffit and Nicholson (1982 in Foley, 1997) found that an increase of 10 points of percentage in the replacement rate caused an effect of a one-week increase in the length of duration. Nevertheless, the probability of exiting unemployment might increase if unemployment payments are no longer distributed. On the contrary, it is not considered an effective decision to implement policies that reduce the number of beneficiaries or even stop distributing unemployment payments (Pellizari, 2006; Kyyrä, Parrotta and Rosholm, 2009; Degen, 2014). Cahuc and Zylberberg (2004) conclude that the optimal level of unemployment

181 The paper is submitted to the International Institute for Private, Commercial and Competition Law in May, 2018.
182 the weekly income percentage that is paid out by unemployment insurance
benefits must be set in such way that creates incentives for job seekers and provides the need to insure unemployed jobseekers against income fluctuations.

According to Erbenova, Sorm and Terell (1998), the social assistance programs support people who are either under social or financial needs. Several early publications concluded that high unemployment rates increase the number of individuals who receive chronical social assistance (Brännström and Stenber, 2007). For evidence, Flister (2015) analyzed the data from the Labor Survey and Income Dynamics in Canada (Saskatchewan) from 1993 to 1994 using an OLS approach and found that an increase of 10 percent in the total social assistance benefits increases duration by 1.4 weeks. Hence, we expect a negative effect of social assistance on exiting unemployment.

Most of the recent publications conclude that there is a double effect of training on unemployment duration. One hand, training programs are considered as layers of control by potential employers (Richardson and Van den Berg, 2002). On the other hand, training might increase the job seeker's reservation wage, and this leads to longer spells (Fougere, Crépon & Ferraci, 2007). Nevertheless, other authors argue that the effect of training on duration might depend on timing. For instance, it is likely that in the short run, e.g., several weeks after the end of the training program, training incentivizes jobseekers to exit unemployment (Fougère, Crépon & Ferraci, 2007; Richardson and Van den Berg, 2002). In addition, McGuinness, O'Connell and Kelly (2014) argue that in the long run, e.g., several months after training, the impact might disappear. Aside from timing, the effect of training on unemployment duration depends also on the nature of training. For instance, Smet (2012) argues that on-job-training raises the probability of employment or reemployment compared to job-search-training.

Levensteijn and Koning (2000) claim that highly skilled job seekers are more preferred by employers. Given that, unemployed jobseekers with high level of education attainment have higher employment chances than individuals with low level of education (Farber, 2005; Nickell, 1997). For instance, Riddell and Song (2011) determined that an increase in education (up to 12 years) will reduce the expected duration by 4 percent. Conversely, it is expected that duration increases for more qualified individuals as long as the employment costs for them are higher (Arendt, Rosholm & Jensen, 2005; Kettuten, 1997). While over-qualification affects duration positively (Lin and Hsu, 2013), Kettuten (1997) finds that education is an insignificant factor. The ambiguousness of the effect is explained in Borjas (2013) highlighting that it is substantial to examine the quality of education rather than its attainment.

There is a lack of specific literature of the effect of gender on unemployment duration (Millard, Dale & Taylor, 1998; Lynch, 1989). Nevertheless, there are empirical studies in which gender is used as an explanatory variable. For instance, Foley (1997) found that the duration for married females is expected to be 10 months longer compared to married males. Additionally, Tansel and Tasçi (2010) in their paper related to the gender effect on unemployment showed that females have a lower probability to find a job compared to males. Regarding age, a lot of authors claim that the closer to the retirement age a jobseeker is, the less attracted becomes the job search process and the higher his chances to re-enter the labor market are (Leuvensteijn & Koning, 2000). While Foley (1997) found that the duration of those in their fifties is anticipated to be longer than the duration of those in their thirties. Nevertheless, there is a low variation of unemployment duration by age and education (Foley, 1997).

3. THE Gompertz Model
In essence, duration models in unemployment explain the time an individual spent in unemployment. In this paper we mostly follow Cameron and Trivedi (2005). Furthermore, we consider only parametric estimators (see Gorenxa, 2018 for non-parametric estimators). The base function in any duration analysis is the hazard function which defines the concept of duration dependence.

\[
\varphi(t) dt = \frac{\Pr[t \leq T \leq t+dt]}{\Pr[T \geq t]} = \frac{dt f(t)}{1-F(t)},
\]

where T is the failure time and \(f(t)\) and \(F(t)\) are the pdf and cdf of the unemployment spell. Intuitively, Eq. (1) reads as the conditional probability that the unemployment duration lies in the \([t, t+dt]\) interval. In addition, the denominator represents the survival function, \(S(t) = 1-F(t)\). Hence, \(\varphi(t) = f(t)/S(t)\).

From the class of parametric estimations in transition analysis, we focus on the proportional hazard (PH) models. The general form of PH models can be written as:

\[
\varphi(t|\mathbf{x}, \boldsymbol{\beta}) = \varphi_0(t) \theta(\mathbf{x}, \boldsymbol{\beta}),
\]

where \(\varphi_0(t)\) denotes the baseline hazard, a function of time only. \(\theta(\mathbf{x}, \boldsymbol{\beta})\) is a function of the covariates in \(\mathbf{x}\). Cameron and Trivedi (2005) argue that the most common specification of \(\theta(\mathbf{x}, \boldsymbol{\beta})\) is the exponential specification. The baseline hazard can follow several distributions including gamma, Gompertz, Weibull, log-normal, log-logistic and exponential. In this paper we consider the Gompertz distribution (\(\varphi_0 \sim \text{Gompertz}(\gamma, \alpha)\)). The hazard pdf can be written as:

\[
f(t) = \gamma e^{\alpha t} e^{-\frac{\gamma}{\alpha}(e^{\alpha t}-1)}
\]

The survival function, \(S(t) = e^{-\frac{\gamma}{\alpha}(e^{\alpha t}-1)}\), and the hazard function, \(\varphi(t) = \gamma \exp(\alpha t)\). To derive the cumulative hazard function, we merely integrate the hazard function, or use the shortcut formula: \(\Phi(t) = -\ln S(t)\).

Hence, the hazard under the Gompertz distribution of \(\varphi_0(t)\) takes the form \(\gamma \exp(\alpha t) \exp(x'\beta)\). To parameterize the model, we apply the logarithmic transformation and write the the parametric Gompertz model as in Eq (3):

\[
\log \varphi(t|\mathbf{x}, \boldsymbol{\beta}) = \log \gamma + \alpha t + x'\beta + \epsilon,
\]

Note that the intercept in the Gompertz model is \(\log \gamma + \alpha t + \beta_0\). The following section presents the data and estimations.

4. DATA AND ESTIMATIONS

The data sample is drawn from the Living Standards Measurement Survey (LSMS) of Albania for the year 2012. LSMS is a multi-purpose households survey. It is considered as one of the main data sources to measure living standards and poverty indicators based on household
consumption. Moreover, the survey provides sufficient and relevant tools to help policy makers in monitoring and developing social programs.

We restrict the sample to only those aged from 19 to 35 years old. In addition, we exclude all individuals with student status in 2012. The survey does not provide a measure of unemployment duration. However, we construct the unemployment spell measure given the information on first employment year and the respondent's age at last year of her studies. We adjust the unemployment spell by subtracting half year to all observations given that on average, the academic year in Albania ends in June. Further, we drop all cases with unemployment spells longer than 19 years given that the minimum working age is 16. The determinants of unemployment duration we consider in this analysis are a set of demographics including age, gender and residence, human capital attributes including education attainment, and occupation dummies. Table (A1) and (A2) (in Appendices) present the sample characteristics.

The following analysis present the parametric estimations of hazard rates. In this work we consider the distributions of the baseline hazard. The Gompertz regression is a proportional hazard model under the Gompertz specification of the baseline hazard. Table (3) presents the estimated model. In addition, the results are not sensitive to unobserved heterogeneity under gamma distribution. Hence, Table (3) reports the estimates of the Gompertz model without accounting for unobserved heterogeneity. The failure time column shows the effect of covariates on the time when record ends. A positive coefficient means lower unemployment duration and vice versa. The hazard rates are interpreted as follows: a hazard rate higher than 1 means a "high failure rate", which in the context of the school to work transition means a shorter duration of first unemployment.

We observe that the unemployment duration of males is shorter than that of females. Moreover, males have 11.5 percent higher hazard rates, i.e., males are more likely to find a job compared to females. Age has a negative effect on unemployment duration. That is, the odds of exiting unemployment decrease with age. In specific, a marginal increase in age, decreases the failure time by 20 percent. In the exponential specification, the failure time does also decelerate by a marginal increase in age. The indicator of urban residence has no significant effect on the failure time. Regarding education, those who have earned a vocational education degree have shorter spells than those who have competed gymnasium or those with lower education attainment. Vocational education graduates have approximately 63 percent higher hazard rates.

<table>
<thead>
<tr>
<th>Covariates (X)</th>
<th>Hazard rate</th>
<th>failure time</th>
<th>S.E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (female)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.153</td>
<td>0.143***</td>
<td>0.060</td>
</tr>
<tr>
<td>Age</td>
<td>0.800</td>
<td>-0.223***</td>
<td>0.008</td>
</tr>
<tr>
<td>Residence (other)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1.058</td>
<td>0.057</td>
<td>0.067</td>
</tr>
<tr>
<td>Schooling (general secondary or less)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocational education</td>
<td>1.628</td>
<td>0.487***</td>
<td>0.144</td>
</tr>
<tr>
<td>Higher education</td>
<td>5.001</td>
<td>1.610***</td>
<td>0.111</td>
</tr>
<tr>
<td>Occupations (elementary occupations)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professionals</td>
<td>1.896</td>
<td>0.640***</td>
<td>0.148</td>
</tr>
</tbody>
</table>
Moreover, the unemployment duration of university graduates is shorter than the duration of general high-school graduates or those with lower education attainment. Regarding occupation dummies, professionals have shorter spells than those employed in elementary occupations. While professionals have 90 percent higher hazard rates than the reference category, those who work as armed forces have 147 percent higher hazard rates than those employed in elementary occupations. In contrast, the estimate reveals insignificance of the armed forces indicator in the exponential specification.

The indicators of managerial and operative positions are insignificant. In addition, technicians and service workers have 73 and 40 percent higher hazard rates compared to elementary occupations employees, respectively. The hazard rates of those employed in clerical jobs are 36.8 percent higher compared to those of elementary occupations. Lastly, the indicators of skilled agriculture and craft and related trades occupations are highly significant determinants of the first period of unemployment. The hazard rates of those employed in skilled agriculture and craft occupations are 75 and 46 percent higher than the hazard rates of elementary occupation workers'.

Intuitively, the results of the Gompertz model fit the reality. Technically, the model choice is in general based on the Akaike's Information Criteria (AIC). The best model would be the one with the lower AIC value. Hence, we prefer the Gompertz specification over the exponential and weibull ones183. However, the Akaike's Information Criteria does not indicate which model provides the best fit to the data. To this extent, further diagnosis emerges. Hence, we can answer what does each model imply regarding the duration dependence (shape of the hazard function) of a spell of unemployment. Figure (1) presents the hazard probability distribution functions184 (pdf) under the exponential, weibull, and log-logistic and Gompertz distributions.

In the case of the exponential distribution the hazard pdf is decreasing. However, we cannot observe the exact spikes and dips in duration dependence. In contrast, the weibull and log-logistic distributions provide a better fit to the data. This is owing to the baseline hazard specification of the weibull model which involves the analysis time. Conversely, the exponential specification does not involve time in the baseline hazard. With respect to the Gompertz distribution, we observe that the hazard pdf does also capture the fluctuations in the

183 see Gorenca (2018) for the estimates under the exponential and weibull distributions, with their respective AIC values.
184 the chosen parameters' values for the exponential, weibull and log-logistic densities are $\gamma=0.01$ and $\alpha=1.5$ base on McCall (1970)
probability of exiting unemployment. However, after the maximum spell observed, the density becomes instantaneously zero. Cameron and Trivedi (2005) argue that the Gompertz model provides a better fit to mortality data. Moreover, the smoothness of the hazard pdf under Gompertz distribution is highly sensitive to the choice of function parameters ($\gamma, \alpha$). In the figure below, $\gamma=0.01$ and $\alpha=0.25$.

**Hazard pdfs under various distributions**

![Hazard pdfs under various distributions](image)

**CONCLUSION**

In this paper we aim to determine the effect of socio-economic factors on unemployment duration. Among the estimated parametric models, we conclude that the Gompertz model has the lowest AIC value. However, the weibull and log-logistic models provide a better fit to the data as the duration dependence (hazard pdf shape) is closer to what we observe in reality. The estimates suggest an important role of factors such as age, education, gender and occupation on unemployment duration. Moreover, males are more likely to find a job than females. The older an individual becomes, the lower the chances to exit from unemployment are. We report insignificance of the urban residence indicator. Professionals, technicians, armed forces, service, craft and related trades, clerical and skilled agriculture workers have higher odds of exiting from unemployment compared to those employed in elementary occupations. The estimated model indicates insignificance of occupations such as managers and operators.

It is worth noting that the socio-economic factors examined in this study do not provide sufficient source of motivation for policymaking. In specific, there is lack of labor market policy indicators such as unemployment benefits and social assistance. Instead, our estimates offer a comparative overview of the job-finding probability among individuals with different socio-economic characteristics. Further examination concerns future research.
REFERENCES


**APPENDICES: A1 & A2**

**TABLE I. SAMPLE CHARACTERISTICS (CATEGORICAL VARIABLES)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Frequency</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>994</td>
<td>32.47</td>
</tr>
<tr>
<td>Females</td>
<td>478</td>
<td>67.53</td>
</tr>
<tr>
<td>Education Attainment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary or less</td>
<td>559</td>
<td>37.98</td>
</tr>
<tr>
<td>General secondary</td>
<td>455</td>
<td>30.91</td>
</tr>
<tr>
<td>Vocational</td>
<td>53</td>
<td>3.60</td>
</tr>
<tr>
<td>Higher education</td>
<td>405</td>
<td>27.51</td>
</tr>
<tr>
<td>Residence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>747</td>
<td>50.75</td>
</tr>
<tr>
<td>Urban</td>
<td>725</td>
<td>49.25</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professionals</td>
<td>244</td>
<td>16.58</td>
</tr>
<tr>
<td>Managers &amp; Admin.</td>
<td>17</td>
<td>1.15</td>
</tr>
<tr>
<td>Service workers</td>
<td>224</td>
<td>15.22</td>
</tr>
<tr>
<td>Clerical</td>
<td>39</td>
<td>2.65</td>
</tr>
<tr>
<td>Operative workers</td>
<td>94</td>
<td>6.39</td>
</tr>
<tr>
<td>Technicians</td>
<td>71</td>
<td>4.82</td>
</tr>
<tr>
<td>Skilled agriculture</td>
<td>430</td>
<td>29.21</td>
</tr>
<tr>
<td>Elementary occupations</td>
<td>103</td>
<td>7.00</td>
</tr>
<tr>
<td>Craft and related trades</td>
<td>210</td>
<td>14.27</td>
</tr>
<tr>
<td>Armed forces</td>
<td>19</td>
<td>1.29</td>
</tr>
</tbody>
</table>

**SAMPLE CHARACTERISTICS (CONTINUOUS VARIABLES)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1,472</td>
<td>28.19</td>
<td>4.32</td>
<td>19</td>
<td>35</td>
</tr>
<tr>
<td>Spell</td>
<td>1,472</td>
<td>6.99</td>
<td>4.53</td>
<td>0.08</td>
<td>18.83</td>
</tr>
</tbody>
</table>