

Phân tích chênh lệch thu nhập theo giới tính ở TP. Hồ Chí Minh bằng hồi quy phân vị

Ngày nhận: 09/07/2013
Ngày nhận lại: 02/10/2013
Ngày duyệt đăng: 31/10/2013
Mã số: 07-13-MA-15

Trần Thị Tuấn Anh
Trường Đại học Kinh tế TP.HCM
anhttt@ueh.edu.vn

Tóm tắt

Quá trình tăng trưởng và phát triển của các quốc gia luôn kéo theo hệ quả về chênh lệch thu nhập, và mức sống trong xã hội. Chênh lệch thu nhập theo giới tính là một trong những vấn đề được quan tâm nghiên cứu nhiều trên thế giới. Bài viết này vận dụng phương pháp hồi quy phân vị trên số liệu về thu nhập khảo sát được trên địa bàn TP.HCM để đo lường mức độ chênh lệch thu nhập theo giới tính ở thành phố này. Kết quả phân tích cho thấy thực sự có chênh lệch thu nhập giữa hai nhóm lao động nam và nữ trên địa bàn TP.HCM. Sự chênh lệch này xảy ra ở cả hai khu vực nội thành và ngoại thành trên tất cả các phân vị được xét. Sự khác biệt về đặc điểm lao động như trình độ học vấn, ngành nghề... thực sự giải thích được sự chênh lệch thu nhập theo giới tính nhưng mức độ giải thích khác nhau ở những phân vị khác nhau.

Từ khóa: Thu nhập, chênh lệch thu nhập, hồi quy phân vị, phân rã chênh lệch.

Abstract

The growth of the nations usually leads to the problem of income disparity. The gender income gap is one of the issues that have been widely studied over the world. In this article, we employed quantile regression techniques to study the gender income gap in HCMC. Moreover, we decomposed the gender income gap into the explained portion by gender differences in characteristics and the unexplained portion. We show that the gender income gap really exists in HCMC, in both urban and rural districts. The explained portion of the gap is about 25%.

Keywords: Gender income gap, quantile regression, decomposition, gender differences in characteristics.

1. Giới thiệu

Các lí thuyết kinh tế cổ điển trước đây cho rằng chênh lệch thu nhập và tiền lương chính là khoản đãi ngộ xứng đáng cho sự khác biệt về năng lực cá nhân hoặc năng suất của người lao động. Những người lao động với năng lực cá nhân cao hơn và năng suất lao động cao thì xứng đáng được hưởng mức thù lao lớn hơn. Tuy nhiên, các lí thuyết này gặp khó khăn trong việc lí giải các mức chênh lệch thu nhập quá lớn xảy ra trong nền kinh tế. Nhiều nghiên cứu áp dụng các phương pháp định lượng cũng chỉ ra rằng sự chênh lệch thu nhập bên cạnh những chênh lệch về đặc điểm và trình độ lao động, dường như bị ảnh hưởng rất lớn bởi các yếu tố không đo lường được năng lực của người lao động như giới tính, dân tộc, tình trạng hôn nhân. Đó có thể là dấu hiệu của sự phân biệt đối xử, của bất bình đẳng trong xã hội. Trong số các phương pháp định lượng để nghiên cứu chênh lệch thu nhập, phương pháp hồi quy phân vị đang ngày càng được sử dụng phổ biến.

Bài viết này sử dụng phương pháp hồi quy phân vị để hồi quy hàm thu nhập và phương pháp của Machado-Mata (2005) để phân tích chênh lệch thu nhập theo giới tính của người lao động ở TP.HCM.

2. Cơ sở lí thuyết

Hồi quy phân vị được Koenker & Bassett đề xuất năm 1978 và được Buchinsky (1994) khởi xướng ứng dụng để xây dựng hàm thu nhập – tiền lương theo trình độ học vấn. Hàm hồi quy thu nhập theo từng phân vị của Buchinsky xây dựng cho thấy tác động của học vấn đến thu nhập tại từng điểm khác nhau của hàm phân phối biến thu nhập. Tiếp theo các nghiên cứu tiên phong của Buchinsky có thể kể đến Chamberlain (1994), Machado & Mata (1999).

Dựa trên phương pháp phân tích chênh lệch tiền lương của Oaxaca – Blinder (1973), Machado & Mata (2005), Melly (2006) xây dựng kĩ thuật phân tích tương tự đối với hàm hồi quy phân vị; từ đó phân tích phần chênh lệch thu nhập tại từng phân vị thành 2 phần chính: phần chênh lệch đã được giải thích gây ra bởi sự khác biệt về các biến độc lập nghiên cứu trong mô hình và phần chênh lệch chưa giải thích, được xem như là phần thể hiện của vấn đề bất bình đẳng về thu nhập.

Nhiều nghiên cứu sau đó cũng đã áp dụng và phát triển các kĩ thuật phân tích chênh lệch thu nhập bằng hồi quy phân vị. Albrecht & cộng sự (2004) đã sử dụng phương pháp phân rã chênh lệch thu nhập theo giới tính bằng hồi quy phân vị ở Hà Lan. Arulampalam & cộng sự (2005) sử dụng phương pháp hồi quy phân vị với số liệu 11 nước châu Âu để phân tích chênh lệch thu nhập theo giới tính và kết luận rằng nữ giới ở đa số các quốc gia đối diện với tình trạng chênh lệch tiền lương rõ rệt so với nam giới.

Machado & Mata (2001) sử dụng số liệu của Bồ Đào Nha và kết luận rằng khoảng cách chênh lệch thu nhập theo giới tính càng lớn khi xét những nhóm phân vị cao, ứng với đối tượng có thu nhập cao. Nhưng mức độ chênh lệch lớn nhất là ở phân vị giữa của hàm phân phối thu nhập. Ngược lại, nghiên cứu của Sakellariou (2004), với số liệu của Philippines, cho thấy nam giới thu nhập cao hơn 50% so với nữ giới ở nhóm phân vị thấp nhất trong khi nữ giới lại có thu nhập thực sự cao hơn nam giới ở những phân vị cao. Mary (2007) hồi quy phân vị với số liệu của vương quốc Anh kết luận mức độ đãi ngộ theo trình độ học vấn ở nước này có sự chênh lệch theo giới tính và mức độ chênh lệch này thay đổi theo thời gian.

Ở VN, vấn đề chênh lệch thu nhập và tiền lương cũng được quan tâm nghiên cứu trong những năm gần đây. Nội dung các nghiên cứu chủ yếu tập trung vào tìm hiểu sự thay đổi của mức độ chênh lệch theo thời gian và nguyên nhân gây ra chênh lệch thu nhập. Các nghiên cứu chênh lệch thu nhập trước đây, trong giai đoạn 1993 – 1998 thường dựa trên hệ số Gini. Hệ số Gini của VN, được tính toán bởi Viện Khoa học xã hội (VASS), tăng từ 0,34 năm 1993 đến 0,35 năm 1998 và đã tăng đến 0,43 năm 2010. Một số nghiên cứu khác cũng dùng hệ số Gini để so sánh bất bình đẳng thu nhập giữa 8 vùng kinh tế ở VN (VASS, 2007) và giữa thành thị - nông thôn (Glewwe, 2000).

Theo Hoang & cộng sự (2001), hồi quy thu nhập ở VN theo đặc điểm lao động giữa khu vực nông thôn – thành thị. Kết quả hồi quy cho thấy thu nhập ở khu vực nông thôn thấp hơn ở khu vực thành thị và chênh lệch này càng lớn trong những năm về sau. Gallup (2004) cũng đưa ra những kết luận tương tự.

Bình T. N & cộng sự (2006) đã sử dụng số liệu của VHLSS từ 1993 đến 1998 để kiểm tra vấn đề bất bình đẳng trong thu nhập và phúc lợi giữa khu vực thành thị và nông thôn ở VN thông qua phương pháp hồi quy phân vị. Kết quả nghiên cứu với số liệu năm 1998, sự chênh lệch của thu nhập được giải thích bởi các biến độc lập xảy ra chủ yếu ở các phân vị thấp trong khi sự khác biệt gây ra chưa được giải thích chủ yếu ở các phân vị cao.

Hung T.P & cộng sự (2006) nghiên cứu chênh lệch thu nhập theo vấn đề dân tộc ở VN. Các tác giả so sánh thu nhập giữa nhóm lao động thuộc các dân tộc Kinh – Hoa so với nhóm còn lại là dân tộc khác. Bài viết đã chỉ ra rằng chênh lệch thu nhập giữa hai nhóm dân tộc lên đến 11%; nhưng chỉ một phần ba của số này là phần chênh lệch được giải thích bởi chênh lệch về đặc điểm lao động trung bình giữa hai nhóm.

Vấn đề chênh lệch tiền lương theo các đặc điểm về học vấn, thành thị, nông thôn... đã được quan tâm nghiên cứu nhưng rất ít các nghiên cứu phân tích chênh lệch theo giới tính. Và chưa có nghiên cứu nào thực hiện phân tích chênh lệch thu nhập theo giới tính tại TP.HCM – một trong những thành phố phát triển sôi động nhất VN.

3. Phương pháp nghiên cứu

Xét hàm hồi quy tuyến tính $Y_i = X_i\beta + U_i$ với Y_i là biến phụ thuộc, X_i là vectơ các biến độc lập; U_i là sai số và hàm hồi quy mẫu ước lượng cho nó $Y_i = X_i\hat{\beta} + e_i$

Theo phương pháp bình phương nhỏ nhất (OLS – Ordinary Least Square), hàm hồi quy mẫu được ước lượng sao cho tổng bình phương sai số là nhỏ nhất, tức là $\hat{\beta} = \arg \min_{\beta} \sum_{i=1}^n (y_i - X_i\beta)^2$

Koenker & Basset (1978) phân tích các nhược điểm của phương pháp OLS. Hai ông kết luận rằng phương pháp OLS chịu sự ràng buộc chặt chẽ của các giả thiết, chịu tác động rất lớn của các quan sát bất thường và không cho thấy cái nhìn toàn diện về hàm phân phối của đại lượng nghiên cứu. Từ đó, Koenker & Basset đề xuất một phương pháp hồi quy mới để khắc phục các nhược điểm này, đó là phương pháp hồi quy phân vị. Thay vì chỉ tiến hành hồi quy để có được hàm trung bình có điều kiện như OLS, hồi quy được thực hiện trên từng phân vị của biến phụ thuộc để cho thấy toàn diện về hàm phân phối của biến phụ thuộc đang được nghiên cứu.

Hàm phân vị tuyến tính có điều kiện của Y theo X ở phân vị $\tau \in (0,1)$ là hàm số $Q_\tau(Y_i) = X_i \hat{\beta}_\tau$ trong đó tham số $\hat{\beta}_\tau$ được chọn sao cho tổng chênh lệch sai số ở phân vị τ nhỏ nhất.

$$\hat{\beta}_\tau = \arg \min_{\beta_\tau} \left(\tau \sum_{y_i \geq X_i \beta_\tau} (y_i - X_i \beta_\tau) + (\tau - 1) \sum_{y_i < X_i \beta_\tau} (y_i - X_i \beta_\tau) \right)$$

Các hàm hồi quy phân vị thường không được sử dụng riêng rẽ ở một phân vị nào mà được tiến hành hồi quy nhiều phân vị một lúc. Trong thời gian đầu, giá trị các hệ số hồi quy phân vị được tính bằng cách giải bài toán quy hoạch tuyến tính với phương pháp đơn hình. Ngày nay, với sự hỗ trợ của máy tính, các tham số của hàm hồi quy phân vị được ước lượng bằng kỹ thuật bootstrap. Kỹ thuật bootstrap được Bradley Efron thuộc Đại học Stanford (Mỹ) phát triển từ cuối thập niên 1970. Đến năm 2002, nó bắt đầu được ứng dụng trong tính toán hệ số của hồi quy phân vị.

Trong hàm hồi quy mà đề tài sử dụng để phân tích, biến phụ thuộc là logarit thu nhập của người lao động trong một tháng. Các biến độc lập gồm các biến liên quan đến đặc điểm nhân khẩu học, năng lực cá nhân và công việc của người lao động.

$$\text{Hàm hồi quy biến thu nhập dạng tuyến tính như sau : } W = X\beta + U, \quad i = \overline{1, n}$$

Trong đó

W_i : Biến log – thu nhập trong 1 tháng của người dân TP.HCM.

X : Các biến độc lập trong mô hình, bao gồm các biến được liệt kê trong Bảng 1.

Hồi quy phân vị biến tiền lương với các biến độc lập trong Bảng 1 được viết như sau :

$$\begin{aligned} Q_\tau(W_i) = & \alpha_{1\tau} + \alpha_{2\tau} \text{GioiTinh}_i + \alpha_{3\tau} \text{DaKetHon}_i + \alpha_{4\tau} \text{Tuoi}_i + \alpha_{5\tau} \text{Tuoi_binhphuong}_i + \alpha_{6\tau} \text{SucKhoe}_i \\ & + \alpha_{7\tau} \text{ThanhThaoNN}_i + \beta_{2\tau} \text{PTCS}_i + \beta_{3\tau} \text{PTTH}_i + \beta_{4\tau} \text{CD}_i + \beta_{5\tau} \text{DH}_i + \beta_{6\tau} \text{SDH}_i \\ & + \beta_{7\tau} \text{ThanhThaoNN}_i + \gamma_{2\tau} \text{CNTT}_i + \gamma_{3\tau} \text{GiaoDuc}_i + \gamma_{4\tau} \text{KinhDoanh}_i + \gamma_{5\tau} \text{Luat}_i \\ & + \gamma_{6\tau} \text{NganhY}_i + \gamma_{7\tau} \text{NgheThuat}_i + \gamma_{8\tau} \text{QuanDoi}_i + \gamma_{9\tau} \text{XayDung}_i \\ & + \gamma_{10\tau} \text{DNNN}_i + \gamma_{11\tau} \text{DNTN}_i + \gamma_{12\tau} \text{LD}_i + \gamma_{13\tau} \text{QuanLy}_i + \varepsilon_i \end{aligned}$$

$Q_\tau(W_i)$: Phân vị τ có điều kiện của biến logarit thu nhập, $\tau \in (0,1)$

Bảng 1. Danh sách các biến phụ thuộc và biến độc lập

Tên biến	Diễn giải
Log_Thunhap	Logarit của biến thu nhập của người dân TP.HCM
Dakethon	Biến giả, nhận giá trị bằng 1 nếu đã kết hôn, bằng 0 nếu chưa kết hôn
Cocoonho	Biến giả, nhận giá trị bằng 1 nếu có con nhỏ còn trong độ tuổi học phổ thông
Suckhoe	Biến giả, bằng 1 nếu người lao động ít bị bệnh trong năm.
Tuoi	Tuổi
Tuoi_Binhphuong	Bình phương của biến tuổi
PTCS	Biến giả, bằng 1 nếu bằng cấp cao nhất là bằng tốt nghiệp phổ thông cơ sở
PTTH	Biến giả, bằng 1 nếu bằng cấp cao nhất là bằng tốt nghiệp phổ thông trung học

CD	Biến giả, bằng 1 nếu bằng cấp cao nhất là bằng cao đẳng hoặc trung cấp nghề
DH	Biến giả, bằng 1 nếu bằng cấp cao nhất là bằng tốt nghiệp đại học
SDH	Biến giả, bằng 1 nếu bằng cấp cao nhất là bằng tốt nghiệp sau đại học
Cntt	Biến giả, bằng 1 nếu làm trong lĩnh vực công nghệ thông tin
Giaoduc	Biến giả, bằng 1 nếu làm trong lĩnh vực giáo dục đào tạo
Kinhdoanh	Biến giả, bằng 1 nếu làm trong lĩnh vực kinh doanh, buôn bán
Luat	Biến giả, bằng 1 nếu làm trong lĩnh vực luật
NganhY	Biến giả, bằng 1 nếu làm trong lĩnh vực y khoa
Nghethuat	Biến giả, bằng 1 nếu làm trong lĩnh vực nghệ thuật
Quandoi	Biến giả, bằng 1 nếu làm trong lĩnh vực quân đội
Xaydung	Biến giả, bằng 1 nếu làm trong lĩnh vực xây dựng
DNNN	Biến giả, bằng 1 nếu làm ở doanh nghiệp nhà nước
DNTN	Biến giả, bằng 1 nếu làm ở doanh nghiệp tư nhân
LD	Biến giả, bằng 1 nếu làm ở doanh nghiệp liên doanh
Nhanvien	Biến giả, bằng 1 nếu công tác ở vị trí một nhân viên
Quanly	Biến giả, bằng 1 nếu công tác ở vị trí một lãnh đạo
ThanhthaoNN	Biến giả, bằng 1 nếu sử dụng thành thạo ít nhất một ngoại ngữ

Giả sử hàm hồi quy của biến thu nhập tại phân vị $\tau \in (0,1)$ được xây dựng cho nhóm lao động nam.

$$Q_{\tau}(W_i^m) = Q(X_i^m, \beta_{\tau}^m) = X_i^m \beta_{\tau}^m$$

Hàm hồi quy của biến thu nhập tại phân vị τ được xây dựng cho nhóm lao động nữ.

$$Q_{\tau}(W_i^f) = X_i^f \beta_{\tau}^f$$

Với

$\beta_{\tau}^m, \beta_{\tau}^f$: hệ số của hàm hồi quy thu nhập của nam/nữ tại phân vị τ

X_i^m, X_i^f : Giá trị biến độc lập ứng với lao động nam/nữ

Chênh lệch thu nhập ở phân vị τ của hai nhóm lao động nam và nữ sẽ là

$$Q_{\tau}(W_i^m) - Q_{\tau}(W_i^f) = [X_i^m \beta_{\tau}^m - X_i^f \beta_{\tau}^f]$$

Áp dụng phương pháp phân tích của Machado – Mata (2005), phân chênh lệch thu nhập trên sẽ được phân tích như sau:

$$Q_{\tau}(W_i^m) - Q_{\tau}(W_i^f) = (X_i^m \beta_{\tau}^m - X_i^m \beta_{\tau}^f) + (X_i^m \beta_{\tau}^f - X_i^f \beta_{\tau}^f) = \underbrace{X_i^m (\beta_{\tau}^m - \beta_{\tau}^f)}_{\Delta D} + \underbrace{(X_i^m - X_i^f) \beta_{\tau}^f}_{\Delta E}$$

Trong đó:

ΔE : Phần chênh lệch của biến log – thu nhập tại phân vị τ được giải thích bởi sự chênh lệch về đặc điểm cá nhân và năng lực giữa hai nhóm lao động nam và nữ.

ΔD : Phần chênh lệch của biến log – thu nhập tại phân vị τ chưa được giải thích, gây ra bởi sự chênh lệch về hệ số hồi quy giữa hàm hồi quy theo phân vị của hai nhóm lao động nam nữ

Kết quả nghiên cứu

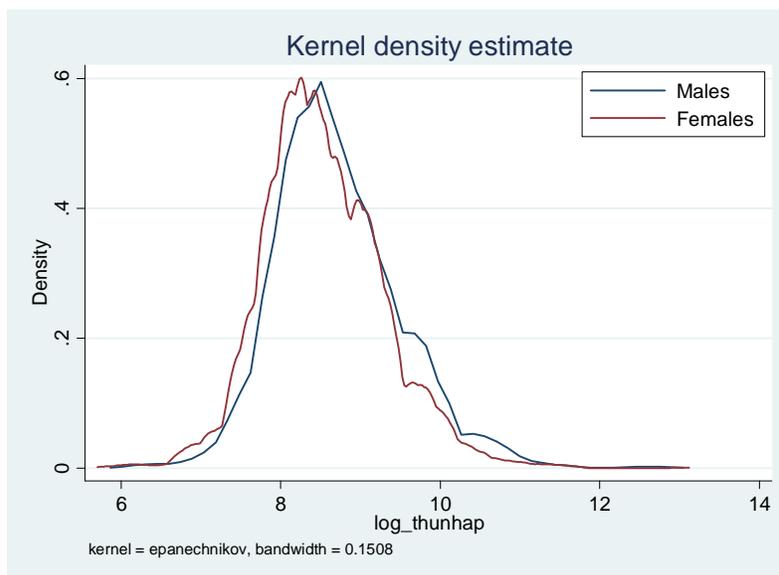
Số liệu được thu thập bằng cách phát phiếu khảo sát để tìm hiểu thu nhập của người dân ở TP.HCM. Mẫu được chọn theo phương pháp phi xác suất có định ngạch. Việc định ngạch được phân bổ theo giới tính và theo 24 quận huyện của TP.HCM. Đợt khảo sát được thực hiện từ tháng 9 đến tháng 12 năm 2011. Sau khi loại bỏ những phiếu khảo sát không trả lời đầy đủ các câu hỏi và không hợp lệ, số lượng phiếu đủ điều kiện đưa vào xử lý gồm 3.987 phiếu chia theo nam và nữ của từng quận huyện được thể hiện ở Bảng 2.

Bảng 2. Số phiếu thu thập theo quận/huyện ở TPHCM

Quận/Huyện	Số phiếu theo giới tính		Tổng số
	Nữ	Nam	
Bình Chánh	16	26	42
Bình Tân	41	73	114
Bình Thạnh	84	82	166
Củ Chi	164	164	328
Cần Giờ	118	105	223
Gò Vấp	105	86	191
Hóc Môn	47	49	96
Nhà Bè	50	59	109
Phú Nhuận	57	60	117
Quận 1	182	194	376
Quận 10	106	143	249
Quận 11	53	56	109
Quận 12	70	91	161
Quận 2	83	102	185
Quận 3	74	114	188
Quận 4	58	87	145
Quận 5	77	88	165
Quận 6	59	98	157
Quận 7	101	125	226
Quận 8	68	78	146
Quận 9	17	18	35
Tân Bình	104	122	226

Thủ Đức	33	38	71
Tân Phú	64	98	162
Tổng cộng	1.831	2.156	3.987

Theo số liệu thu nhập được, hàm mật độ phân phối biến logarit của thu nhập được biểu diễn ở Hình 1.



Hình 1. Hàm mật độ thu nhập theo giới tính nam và nữ

Nguồn: Tác giả tự tính toán từ số liệu thu thập được.

Hình 1 cho thấy hàm mật độ của nhóm lao động nam và nữ có hình dạng và độ cao khá giống nhau nhưng nhóm lao động nữ có xu hướng xiên phải (right - skewed) nhiều hơn một chút.

Kết quả hồi quy biến logarit thu nhập được thực hiện theo phương pháp OLS và hồi quy phân vị trên các phân vị 0,1; 0,25; 0,5; 0,75 và 0,9 được thể hiện trên Bảng 3.

Bảng 3. Kết quả hồi quy OLS và hồi quy phân vị biến log-thu nhập

Các biến độc lập	OLS	Hồi quy phân vị				
		0,10	0,25	0,50	0,75	0,90
GioiTinh	0,112*** [5,639]	0,0797** [2,610]	0,0905*** [4,235]	0,101*** [5,118]	0,121*** [4,529]	0,175*** [4,514]
DaKetHon	0,0803** [3,172]	0,0926** [2,654]	0,0599* [2,278]	0,0795** [3,166]	0,0854* [2,488]	0,0738 [1,539]
Tuoi	0,0833*** [12,651]	0,0760*** [8,477]	0,0872*** [12,243]	0,0879*** [13,491]	0,0837*** [9,340]	0,0922*** [6,918]
tuoi_binhphuong	-0,000954 *** [-11,960]	-0,000937 *** [-8,842]	-0,00106 *** [-12,342]	-0,00104 *** [-13,209]	-0,000938 *** [-8,592]	-0,000990 *** [-6,029]

SucKhoe	0,103** [2,775]	0,154** [2,823]	0,0524 [1,313]	0,0863* [2,366]	0,126* [2,552]	0,142* [2,011]
LamONoiThanh	0,120*** [6,184]	0,122*** [4,201]	0,102*** [4,872]	0,127*** [6,586]	0,111*** [4,256]	0,0800* [2,128]
ThanhThaoNN	0,253*** [9,845]	0,228*** [5,621]	0,239*** [8,573]	0,219*** [8,586]	0,265*** [7,614]	0,332*** [6,507]
PTCS	0,162*** [3,732]	0,188** [2,844]	0,179*** [3,844]	0,169*** [3,934]	0,0516 [0,887]	0,142 [1,699]
PTTH	0,340*** [8,369]	0,334*** [5,559]	0,343*** [7,967]	0,297*** [7,392]	0,194*** [3,525]	0,343*** [4,259]
CD	0,385*** [8,755]	0,409*** [6,323]	0,429*** [9,246]	0,365*** [8,388]	0,244*** [4,042]	0,341*** [3,780]
DH	0,744*** [17,244]	0,770*** [11,881]	0,800*** [17,526]	0,729*** [17,080]	0,645*** [10,825]	0,722*** [8,163]
SDH	1,172*** [19,412]	1,202*** [13,355]	1,225*** [19,137]	1,177*** [19,678]	1,019*** [12,386]	1,080*** [8,745]
CNTT	0,0722 [1,644]	0,0751 [1,165]	0,0331 [0,703]	0,0528 [1,214]	0,0735 [1,249]	0,0114 [0,133]
GiaoDuc	-0,110** [-2,705]	-0,137* [-2,156]	-0,159*** [-3,631]	-0,181*** [-4,483]	-0,159** [-2,909]	-0,0574 [-0,718]
KinhDoanh	0,102*** [4,438]	0,109** [3,081]	0,0676** [2,724]	0,0681** [2,986]	0,0898** [2,933]	0,141** [3,123]
Luat	-0,216* [-2,121]	-0,223 [-1,458]	-0,19 [-1,743]	-0,254* [-2,538]	-0,249 [-1,884]	-0,142 [-0,906]
NganhY	0,101* [2,221]	-0,0785 [-1,137]	-0,0544 [-1,110]	0,023 [0,510]	0,177** [2,900]	0,343*** [3,861]
NgheThuat	0,320*** [4,214]	0,285* [2,464]	0,205* [2,522]	0,306*** [4,119]	0,372*** [3,680]	0,422** [2,876]
QuanDoi	-0,206** [-2,887]	-0,066 [-0,643]	-0,0899 [-1,186]	-0,203** [-2,882]	-0,336*** [-3,497]	-0,290* [-2,086]
XayDung	0,190*** [4,602]	0,201** [3,179]	0,153*** [3,393]	0,170*** [4,164]	0,172** [3,134]	0,162* [2,065]
DNNN	-0,0197 [-0,668]	0,129** [3,005]	0,0355 [1,121]	-0,0145 [-0,496]	-0,026 [-0,670]	-0,130* [-2,342]
DNTN	0,131***	0,231***	0,141***	0,0863***	0,0727*	0,0173

	[5,402]	[6,472]	[5,503]	[3,602]	[2,252]	[0,359]
LD	0,265***	0,338***	0,279***	0,246***	0,208***	0,156*
	[7,485]	[6,420]	[7,299]	[6,997]	[4,435]	[2,343]
QuanLy	0,405***	0,228***	0,276***	0,389***	0,521***	0,588***
	[17,173]	[6,118]	[10,736]	[16,631]	[16,481]	[12,931]
_cons	5,957***	5,473***	5,709***	5,972***	6,335***	6,327***
	[46,106]	[30,597]	[40,683]	[46,749]	[36,228]	[23,795]

t - statistics trong []

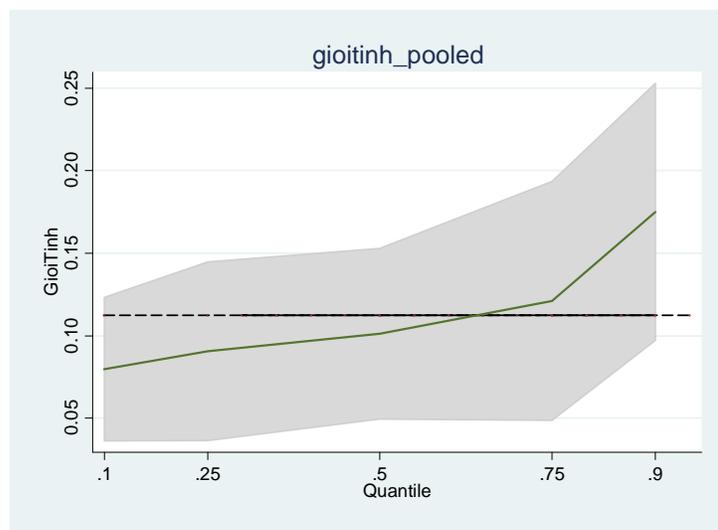
*, **, ***: Có ý nghĩa thống kê với 5%, 1% và 0,1%

Nguồn: Tác giả tự tính toán theo số liệu thu thập được

Kết quả hồi quy OLS của biến log-thu nhập trên toàn bộ mẫu số liệu cho thấy thực sự có chênh lệch thu nhập giữa lao động nam và nữ, mức chênh lệch thu nhập giữa hai nhóm này lên đến 11%. Ngoài ra kết quả ở Bảng 2 hàm ý có sự chênh lệch thu nhập theo giới tính, tình trạng hôn nhân, sử dụng ngoại ngữ thành thạo. Bằng cấp cũng có tác động rất lớn đến thu nhập và người lao động làm việc ở những ngành nghề khác nhau, loại hình doanh nghiệp và vị trí công việc khác nhau thì mức thu nhập cũng rất khác nhau.

Tiến hành kiểm định Breusch-Pagan /Cook-Weisberg về hiện tượng phương sai thay đổi trên mẫu số liệu. Kết quả kiểm định cho thấy với mức ý nghĩa 1%, hàm hồi quy OLS thực hiện như trên có xảy ra hiện tượng phương sai thay đổi. Theo Deaton (1997), khi có hiện tượng phương sai không thuần nhất xảy ra thì kết quả hồi quy phân vị thu được sẽ có những tính chất ưu việt hơn so với hồi quy bằng phương pháp bình phương nhỏ nhất OLS.

Kết quả hồi quy phân vị ở Bảng 3 chỉ ra chênh lệch thu nhập theo giới tính giữa nhóm lao động nam và nữ xảy ra ở tất cả các phân vị được xét. Ở những phân vị càng cao thì mức độ chênh lệch càng cao, xu hướng thể hiện rõ ở Hình 2.

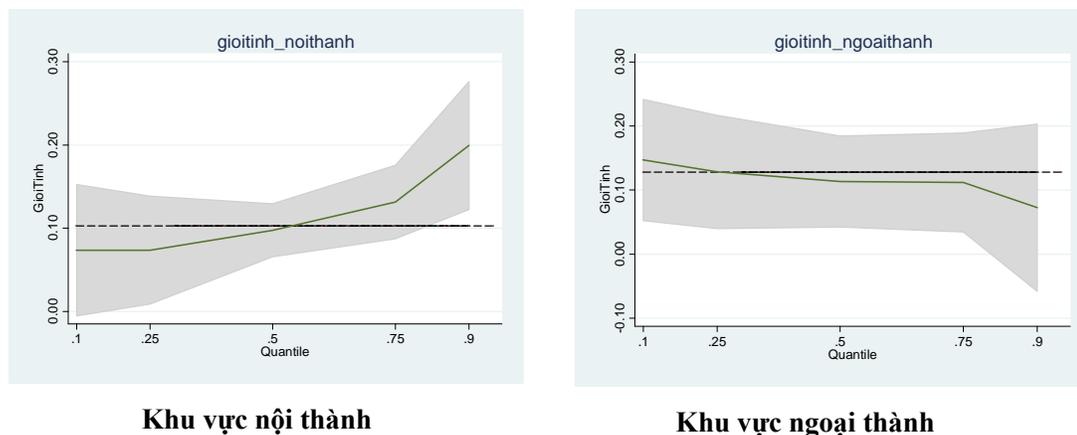


Hình 2. Chênh lệch thu nhập giữa nam và nữ

Nguồn: Tác giả tự tính toán theo số liệu thu thập được

Nhìn chung, nhóm lao động nam có thu nhập cao hơn so với nhóm lao động nữ nhưng ở những phân vị khác nhau thì mức độ chênh lệch khác nhau. Cụ thể, ở phân vị càng cao thì mức chênh lệch càng tăng. Ở phân vị 0,1, thu nhập trung bình của nhóm của lao động nam cao hơn nhóm lao động nữ gần 8%; con số này ở phân vị 0,25 là 9%; ở phân vị 0,5 là 10%; ở phân vị 0,75 là 12% và mức chênh lệch cao nhất ở phân vị 0,9 là 17,5%.

Để làm rõ hơn mức độ chênh lệch thu nhập theo giới tính, bài viết tiến hành hồi quy phân vị biến log-thu nhập ở từng khu vực nội thành và ngoại thành. Kết quả thể hiện ở Bảng 3 và Hình 3.



Hình 3. Chênh lệch thu nhập theo giới tính ở nội thành và ngoại thành

Nguồn: Tác giả tự tính toán theo số liệu thu thập được

Kết quả cho thấy chênh lệch thu nhập ở khu vực nội thành diễn tiến rất giống với chênh lệch thu nhập chung theo giới tính ở thành phố, nghĩa là tại những phân vị càng cao thì mức chênh lệch càng lớn. Trong khi diễn tiến xảy ra hoàn toàn ngược lại ở khu vực ngoại thành.

Bảng 4. Chênh lệch thu nhập giữa nam và nữ theo khu vực nội thành và ngoại thành

Phân vị	0,10	0,25	0,50	0,75	0,90
Nội thành	0,0735 (1,80)	0,0736** (2,12)	0,0974*** (4,15)	0,1314*** (5,27)	0,1994*** (4,63)
Ngoại thành	0,1471** (2,95)	0,1284*** (3,63)	0,1134*** (5,22)	0,1120** (2,95)	0,0725 (0,84)

t-statistic trong ngoặc đơn

*, **, *** : Có ý nghĩa với 5%, 1% và 0,1%

Nguồn: Tác giả tự tính toán từ số liệu thu thập được

Đối với nhóm lao động làm việc ở khu vực nội thành, ở phân vị 0,9: thu nhập của nhóm lao động nam cao hơn nhóm lao động nữ lên đến 20%. Chênh lệch này là 13% ở phân vị 0,75 và giảm xuống 7% ở phân vị 0,25. Ở mức phân vị 0,1: sự chênh lệch thu nhập giữa hai nhóm lao động nam và nữ không có ý nghĩa thống kê.

Khi xét hàm thu nhập ở khu vực ngoại thành, mức chênh lệch cao nhất xảy ra ở nhóm phân vị 0,1 và giảm dần xuống thấp nhất ở mức phân vị 0,9. Theo đó, thu nhập của lao động nam cao

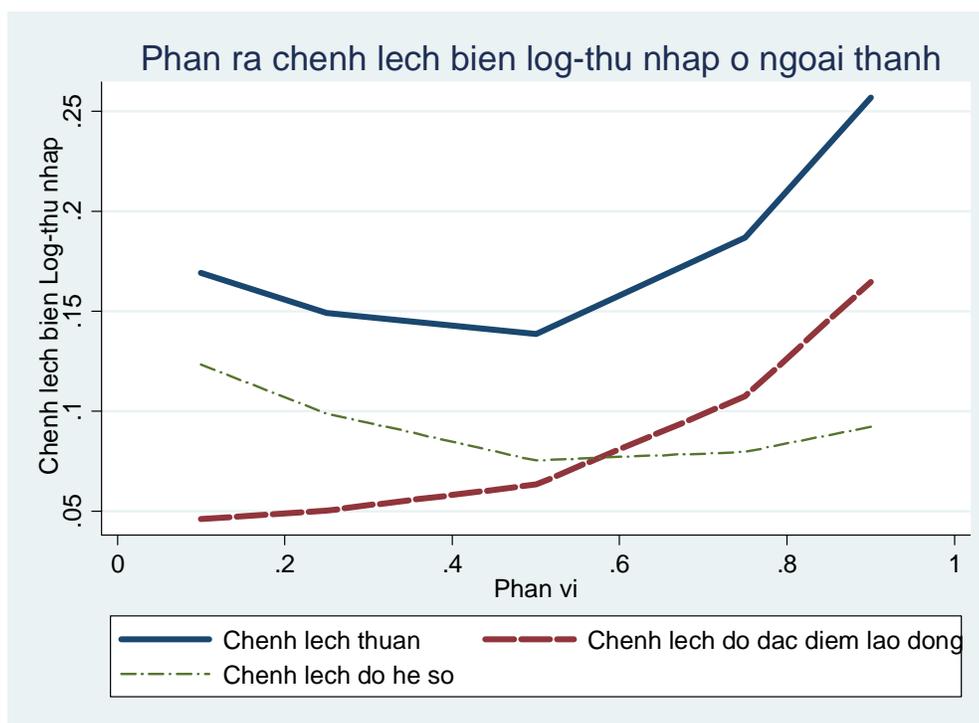
hơn lao động nữ là 14,7% ở phân vị 0,1; con số đó ở phân vị 0,25 là 12,8%; ở phân vị 0,5 là 11,3% và phân vị 0,75 là 11,2%. Sự chênh lệch này ở phân vị 0,9 không có ý nghĩa thống kê.

Theo phương pháp phân rã chênh lệch của Machado & Mata (2005), sự chênh lệch thu nhập giữa nhóm lao động nam và nhóm lao động nữ được chia thành hai phần: phần chênh lệch được giải thích và phần chênh lệch chưa được giải thích. Phần chênh lệch được giải thích là phần chênh lệch gây ra bởi sự khác nhau về đặc điểm nhân khẩu, bằng cấp hoặc các đặc điểm liên quan đến công việc giữa hai nhóm lao động. Trong khi đó, phần chênh lệch chưa được giải thích là phần chênh lệch gây ra bởi sự khác nhau trong hệ số hồi quy khi nghiên cứu hồi quy thu nhập giữa hai nhóm lao động. Trong nhiều nghiên cứu, sự chênh lệch trong chính sách đãi ngộ giữa hai nhóm lao động nam và nữ là dấu hiệu thể hiện của sự bất bình đẳng, của sự phân biệt đối xử giữa nam và nữ (Gunawardena D., 2006; Asplund R & Napari S., 2011...)

Bảng 5. Phân tích chênh lệch thu nhập theo giới tính tại TP.HCM

Thành phần	Hệ số	Sai số chuẩn	T	p-value
Phân vị 0,1				
Chênh lệch thuần	0,159821	0,020624	7,75	0,000
Chênh lệch do đặc điểm lao động	0,043824	0,025308	1,73	0,040
Chênh lệch do hệ số	0,115998	0,021377	5,43	0,000
Phân vị 0,25				
Chênh lệch thuần	0,151194	0,015398	9,82	0,000
Chênh lệch do đặc điểm lao động	0,051894	0,019802	2,62	0,003
Chênh lệch do hệ số	0,0993	0,017282	5,75	0,000
Phân vị 0,5				
Chênh lệch thuần	0,140701	0,020666	6,81	0,000
Chênh lệch do đặc điểm lao động	0,062608	0,022847	2,74	0,000
Chênh lệch do hệ số	0,078093	0,016496	4,73	0,000
Phân vị 0,75				
Chênh lệch thuần	0,186417	0,02378	7,84	0,000
Chênh lệch do đặc điểm lao động	0,105952	0,034992	3,03	0,000
Chênh lệch do hệ số	0,080465	0,022387	3,59	0,000
Phân vị 0,90				
Chênh lệch thuần	0,25964	0,025381	10,23	0,000
Chênh lệch do đặc điểm lao động	0,163256	0,046631	3,50	0,000
Chênh lệch do hệ số	0,096384	0,03388	2,84	0,004

Nguồn: Tác giả tự tính toán từ số liệu thu thập được



Hình 4. Phân tích chênh lệch thu nhập theo giới tính tại TP.HCM

Nguồn: Tác giả tự tính toán theo số liệu thu thập được

Xét trên toàn bộ mẫu số liệu, phần chênh lệch thu nhập được giải thích ở các phân vị khác nhau là khác nhau. Ở phân vị 0,1; mức chênh lệch biến log-thu nhập của lao động nam và nữ lên đến 0,15; trong đó 27,42% mức chênh lệch này được giải thích là do sự chênh lệch của các biến độc lập trong mô hình; còn lại 72,58% là do sự chênh lệch về hệ số hồi quy, là phần chênh lệch do khác biệt trong chính sách đãi ngộ giữa lao động nam và nữ với giả định hai nhóm lao động này có cùng đặc điểm lao động.

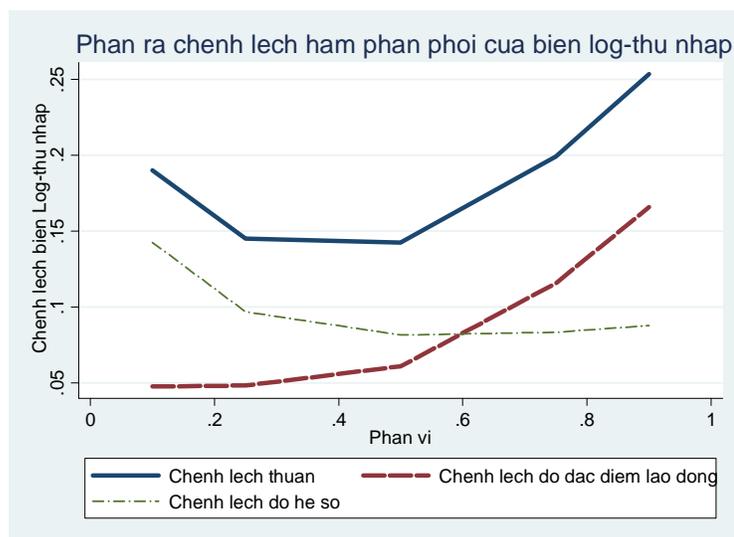
Kết quả phân tích chênh lệch theo từng phân vị trong Bảng 5 cho thấy, ở các phân vị càng cao thì chênh lệch thu nhập cũng càng cao, đồng thời phần chênh lệch thu nhập được giải thích cũng càng nhiều. Cụ thể, ở phân vị 0,25; phần chênh lệch được giải thích là 34,32%; con số này là 44,5% ở phân vị 0,5; ở phân vị 0,77 là 56,84% và đặc biệt ở phân vị 0,9 thì phần chênh lệch chưa được giải thích chỉ còn 37,12%.

Bảng 6. Phân tích chênh lệch thu nhập theo giới tính tại TP.HCM, nhóm làm việc ở nội thành

Thành phần	Hệ số	Sai số chuẩn	T	p-value
Phân vị 0,1				
Chênh lệch thuần	0,190296	0,021684	8,78	0,000
Chênh lệch do đặc điểm lao động	0,047664	0,031759	1,5	0,066
Chênh lệch do hệ số	0,142632	0,025898	5,51	0,000
Phân vị 0,25				

Chênh lệch thuần	0,14518	0,020099	7,22	0,000
Chênh lệch do đặc điểm lao động	0,048278	0,030313	1,59	0,019
Chênh lệch do hệ số	0,096902	0,020666	4,69	0,000
Phân vị 0,5				
Chênh lệch thuần	0,142624	0,025353	5,63	0,000
Chênh lệch do đặc điểm lao động	0,061028	0,027079	2,25	0,007
Chênh lệch do hệ số	0,081596	0,022446	3,64	0,000
Phân vị 0,75				
Chênh lệch thuần	0,199139	0,028711	6,94	0,000
Chênh lệch do đặc điểm lao động	0,115761	0,034593	3,35	0,000
Chênh lệch do hệ số	0,083378	0,028407	2,94	0,003
Phân vị 0,9				
Chênh lệch thuần	0,253606	0,036226	7	0,000
Chênh lệch do đặc điểm lao động	0,165771	0,052403	3,16	0,000
Chênh lệch do hệ số	0,087835	0,035036	2,51	0,012

Nguồn: Tác giả tự tính toán từ số liệu thu thập được



Hình 5. Phân tích chênh lệch thu nhập theo giới tính tại TPHCM ở khu vực nội thành

Nguồn: Tác giả tự tính toán từ số liệu thu thập được

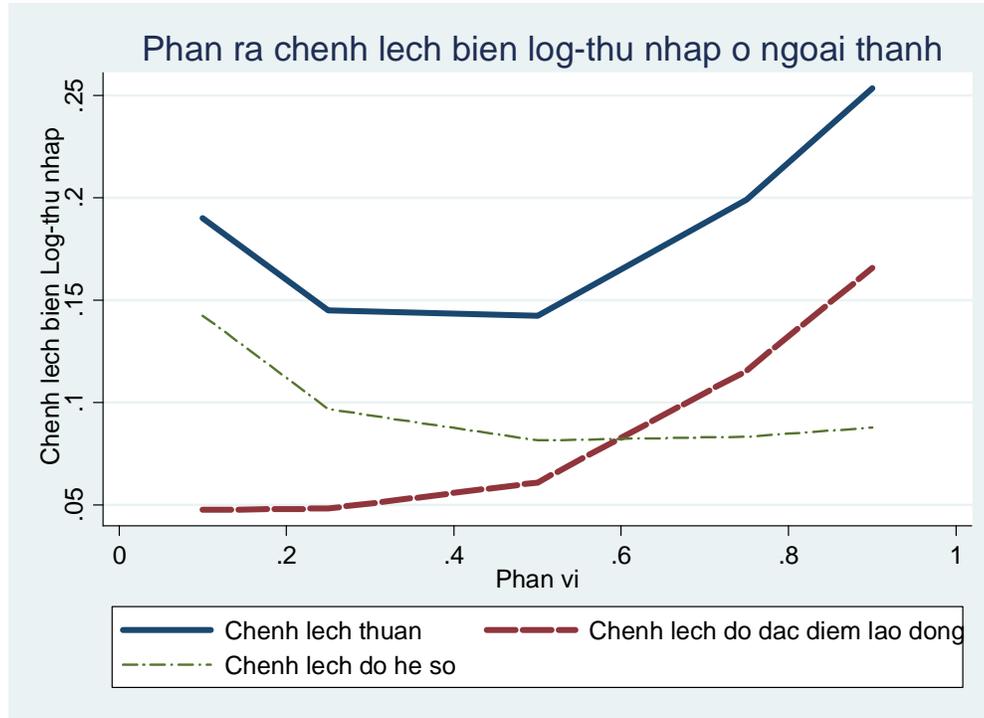
Khi xét phân rã chênh lệch thu nhập của người lao động nam và nữ cùng làm việc ở nội thành, diễn tiến khá giống với trường hợp phân tích chênh lệch trên toàn bộ mẫu số liệu. Điều này có thể quan sát được về sự tương đồng trong hình dạng đường đứt quãng in đậm trong Hình 5. Cụ thể, khi xét nhóm làm việc ở nội thành, ở phân vị 0,1; phần chênh lệch được giải thích bởi các biến độc lập trong mô hình là 25,05% trong khi con số này ở phân vị cao nhất – phân vị 0,9 – lên đến 65,37%. Chênh lệch thu nhập ở nhóm làm việc trong nội thành cũng có hình dạng chữ

U như khi phân tích trên nhóm làm việc ngoại thành (Hình 5) và toàn bộ mẫu số liệu (Hình 4). Nhưng chữ U trong nhóm làm việc ở nội thành này rõ rệt hơn.

Bảng 7. Phân tích chênh lệch thu nhập theo giới tính tại TP.HCM ở khu vực ngoại thành

Thành phần	Hệ số	Sai số chuẩn	t	p-value
Phân vị 0,1				
Chênh lệch thuần	0,124942	0,037169	3,36	0,001
Chênh lệch do đặc điểm lao động	0,011726	0,044843	0,26	0,756
Chênh lệch do hệ số	0,113215	0,037783	3,00	0,003
Phân vị 0,25				
Chênh lệch thuần	0,118488	0,023422	5,06	0,000
Chênh lệch do đặc điểm lao động	0,024201	0,026085	0,93	0,410
Chênh lệch do hệ số	0,094287	0,02937	3,21	0,001
Phân vị 0,5				
Chênh lệch thuần	0,150033	0,026768	5,60	0,000
Chênh lệch do đặc điểm lao động	0,055087	0,030742	1,79	0,038
Chênh lệch do hệ số	0,094946	0,026561	3,57	0,000
Phân vị 0,75				
Chênh lệch thuần	0,173872	0,039995	4,35	0,000
Chênh lệch do đặc điểm lao động	0,089606	0,049728	1,80	0,010
Chênh lệch do hệ số	0,084267	0,034953	2,41	0,016
Phân vị 0,9				
Chênh lệch thuần	0,255556	0,063253	4,04	0,000
Chênh lệch do đặc điểm lao động	0,154377	0,092133	1,68	0,002
Chênh lệch do hệ số	0,101178	0,048778	2,07	0,038

Nguồn: Tác giả tự tính toán từ số liệu thu thập được



Hình 6. Phân tích chênh lệch thu nhập theo giới tính tại TP.HCM khu vực ngoại thành

Nguồn: Tác giả tự tính toán theo số liệu thu thập được

Khi phân tích chênh lệch giữa lao động nam và nữ làm việc ở khu vực ngoại thành không có dạng chữ U rõ rệt như khu vực nội thành nhưng chiều hướng biến đổi khá tương đồng. Chênh lệch cao ở những phân vị đuôi và thấp dần khi tiến vào những phân vị giữa. Mức độ được giải thích ở những phân vị càng cao thì càng nhiều. Cụ thể, ở phân vị 0,1; phần được giải thích chỉ có 9,39%; ở phân vị 0,5 phần chênh lệch được giải thích là 20,42%; ở phân vị 0,5 phần được giải thích là 36,72%; con số này ở phân vị 0,75 là 51,54% trong khi con số cao nhất ở phân vị 0,9 là 60,41%. Phần chênh lệch được giải thích bởi đặc điểm lao động ở phân vị 0,1 của cả hai nhóm lao động nội thành và ngoại thành đều không có ý nghĩa thống kê.

5. Kết luận và kiến nghị

5.1. Kết luận

Phân tích trên mẫu số liệu về thu nhập thu thập được trên địa bàn TP.HCM cho thấy thực sự có chênh lệch thu nhập giữa lao động nam và nữ. Lao động nam thực sự có thu nhập cao hơn lao động nữ. Sự chênh lệch thu nhập này xảy ra khi kiểm định trên tất cả các phân vị của hàm phân phối biến thu nhập. Phân vị càng cao thì mức chênh lệch càng lớn, chênh lệch thu nhập giữa lao động nam và nữ có ý nghĩa thống kê trong nhóm làm việc ở nội thành ở các phân vị từ 0,25 đến phân vị 0,9. Trong khi chênh lệch này xảy ra ở tất cả các phân vị của lao động nội thành thì ở nhóm lao động ngoại thành, chênh lệch thu nhập giữa nam và nữ ở những phân vị càng cao thì càng thấp và ngược lại.

Xét trên toàn bộ mẫu số liệu thu thập ở TP.HCM, sự chênh lệch về đặc điểm lao động chỉ giải thích được 27,42% sự chênh lệch thu nhập ở phân vị 0,1. Và ở những phân vị cao hơn, mức độ

tham gia giải thích của yếu tố đặc điểm cũng nhiều hơn. Nhìn chung, mức độ giải thích của các yếu tố tham gia vào mô hình đối với chênh lệch tiền lương ở ngoại thành cao hơn ở nội thành.

5.2. Một số đề xuất

Sự chênh lệch thu nhập theo giới tính được gây ra bởi rất nhiều nguyên nhân. Một trong số đó chính là chênh lệch về các đặc điểm lao động như trình độ học vấn, mức độ thành thạo ngoại ngữ, sức khỏe. Để thu hẹp khoảng cách thu nhập, nữ giới cần được tạo điều kiện học tập, thăng tiến hơn nữa nhằm giảm chênh lệch trình độ với lao động nam giới. Bên cạnh đó, phụ nữ còn phải đảm đương cả những công việc không được trả công tại gia đình nhiều hơn nam giới. Người phụ nữ còn bị gián đoạn công việc bởi thời gian mang thai, sinh con và nuôi dưỡng con cái. Những điều này là nguyên nhân gián tiếp tác động làm cho thu nhập nữ giới thấp hơn nam giới. Do vậy, việc bình đẳng giữa nam và nữ không phải chỉ được thực hiện tại nơi làm việc mà còn ở trong chính mỗi gia đình của người lao động.

Bên cạnh sự chênh lệch thu nhập gây ra bởi chênh lệch đặc điểm lao động, kết quả nghiên cứu cũng cho thấy phần lớn sự chênh lệch thu nhập gây ra bởi sự chênh lệch hệ số lao động, thể hiện của cơ chế đãi ngộ khác nhau dành cho lao động nam và nữ. Theo Arulampalam & cộng sự (2005), đó có thể là dấu hiệu của tình trạng bất bình đẳng theo giới tính. Ở VN, không thể phủ nhận rằng định kiến phân biệt giữa nam nữ vẫn còn tồn tại. Do vậy, cần kiểm soát lại chính sách đãi ngộ ở những đơn vị sử dụng lao động, kiểm soát lại cơ chế trả lương, trả công, công tác cán bộ của các tổ chức nhằm hạn chế tình trạng phân biệt đối xử theo giới tính. Các chính sách và giải pháp cần hướng tới việc hạn chế bất bình đẳng theo giới tính, tạo cơ hội để nữ giới vươn lên tiếp cận với mức chung của xã hội, chứ không hướng đến cào bằng thu nhập giữa lao động nam và nữ. Khoảng cách chênh lệch thu nhập này không thể giải quyết một cách nóng vội mà đòi hỏi một quá trình nỗ lực không ngừng từ mỗi cá nhân, gia đình, doanh nghiệp và của cả xã hội.

5.3. Hạn chế của nghiên cứu

Bài viết đã phản ánh được một phần thực trạng chênh lệch về thu nhập theo giới tính ở TP.HCM. Tuy nhiên vẫn còn có những hạn chế nhất định. Mẫu số liệu thu thập gần 4.000 quan sát với cách chọn mẫu phi xác suất có thể chưa thực sự đại diện tốt cho tổng thể cần nghiên cứu. Việc thu thập số liệu và phân tích của nghiên cứu được thực hiện trong một năm 2011, chỉ phản ánh được một phần thực trạng chênh lệch thu nhập theo giới tính trong năm đó mà chưa phản ánh được xu hướng thay đổi chênh lệch thu nhập theo giới tính tại TP.HCM theo thời gian. Bên cạnh đó, từ lúc thu thập số liệu cho đến khi nhập và xử lý số liệu mất một khoảng thời gian khá dài, làm cho kết quả không còn mang tính thời sự, hiện trạng tiền lương hiện nay có thể đã thay đổi so với thời điểm bắt đầu tiến hành nghiên cứu. Với những hạn chế như vậy, nghiên cứu có thể được đầu tư mở rộng theo không gian hoặc thời gian để nâng cao tính tổng quát của kết quả thu được. Việc mở rộng theo không gian để nghiên cứu cho một phạm vi lớn hơn như ở Nam Bộ hoặc ở VN. Bài viết cũng có thể mở rộng nghiên cứu theo thời gian, tiến hành thu thập số liệu hơn trong nhiều thời điểm để phân tích, so sánh, làm rõ xu hướng thay đổi chênh lệch thu nhập theo giới tính ở TP.HCM nói riêng và ở VN nói chung ■

Tài liệu tham khảo

- Albrecht J, Vuuren A. V., & Vroman S, (2004), “Decomposing the Gender Wage Gap in the Netherlands with Sample Selection Adjustments”, *Tinbergen Institute Discussion Papers* 04-123/3, Tinbergen Institute.
- Arulampalam, W., Booth, A.L. and Bryan, M.L. (2005), “Is There a Glass Ceiling over Europe? An Exploration of Asymmetries in the Gender Pay Gap Across the Wages Distribution”, *Revised version of IZA Discussion Paper No. 1373*.
- Asplund R.& Napari S., (2011), *Intangibles and the Gender Wage Gap: An Analysis of Gender Wage Gaps Across Occupations in the Finnish Private Sector*, Discussion Papers 1243, The Research Institute of the Finnish Economy.
- Binh N. T., Albrecht, J. W., Vroman, S. and Westbrook, M. D. (2006), “A Quan-Tile Regression Decomposition of Urban-Rural Inequality in Vietnam”, *Journal của Development Economics* 83, 466-490.
- Blinder, A. S. (1973), “Wage Discrimination: Reduced Form and Structural Estimates”, *The Journal of Human Resources* 8: 436-455.
- Buchinsky, M., (1994), “Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression”, *Econometrica* 62,405-458.
- Chamberlain, G., (1994), “Quantile Regression, Censoring and the Structure of Wages” in: C.Sims and J.J.Laffont, eds., *Proceedings of the Sixth World Congress of the Econometrics Society*, (New York: Cambridge University Press), 171-209.
- Deaton, A. (1997), *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*, John Hopkins University Press.
- Gallup J.L. (2004), “The Wage Labour Market and Inequality in Vietnam”, In: Glewwe P, Agarwal N, Dollar D(eds) *Economic Growth, Poverty, and Household Welfare in Vietnam*, The World Bank, Washington DC pp 53-93.
- Glewwe, P., M. Gragnolati, and H. Zaman (2000), “Who Gained from Vietnam’s Boom in the 1990s? An Analysis of Poverty and Inequality Trends”, *World Bank Working Paper 2275*, Washington, D.C.
- Gunawardena, D. (2006), “Exploring Gender Wage Gaps in Sri Lanka: A Quantile Regression”, Approach Paper Presented During the 5th PEP Research Network General Meeting, June 18-22, Addis Ababa, Ethiopia
- Hoang, K., Baulch, B., Le, D., Nguyen, D., Ngo, G., and Nguyen, K., (2001), “Determinants of Earned Income”, in Haughton, J., Haughton, D., and Nguyen, P., eds., *Living Standards During an Economic Boom: The Case of Vietnam*, UNDP and Statistical Publishing House, Hanoi.
- Hung , P. T. and B. Reilly, B. (2006), *The Gender Gap in Vietnam, 1993-2002: A Quantile Regression Approach*, *Poverty Research Unit at Sussex*, Working Paper 34.
- Koenker, R. and G. Bassett (1978), “Regression Quantiles”, *Econometrica* , Vol. 46, pp. 33-50.
- Machado J, and Mata J (2005), “Counter Factual Decomposition of Changes in Wage Distributions using Quantile Regression”, *Journal of Applied Econometrics* 20:445-46
- Mary C. ,(2007), “Closing the Global Gender Pay Gap: Securing Justice for Women’s Work 28 Comp”, *Labor law & Pol’Y Journal*, 219
- Melly, B. (2006), *Estimation of Counterfactual Distributions Using Quantile Regression*, *Mimeo*, Swiss Institute for International Economics and Applied Economic Research (SIAW), University of St. Gallen.
- Rica, S., Dolado, J. and V. Llorens (2005), “Ceilings and Floors: Gender Wage Gaps by Education in Spain”, *IZA DP no. 1483*.
- Sakellariou C. (2004), “Gender-Earnings Differentials Using Quantile Regressions”, *Journal of Labor Research*, vol. 25 (3).
- VASS (Viet Nam Academy of Social Science) (2007), *Poverty Update*, VASS, Hanoi.